This new Dover edition,
first published in 1955, is an unabridged
and unaltered republication of the book
first published in 1940 under the title
"The Philosophy of Peirce: Selected Writings."
It is published through special arrangement with
Routledge and Kegan Paul Ltd.

International Standard Book Number: 0-486-20217-8

Library of Congress Catalog Card Number: 56-13549

This edition is manufactured in the
United States of America and is for sale only in the
United States of America.

Dover Publications, Inc.
180 Varick Street
New York, N. Y. 10014

PHILOSOPHICAL WRITINGS OF PEIRCE

Selected and Edited
With an Introduction by
JUSTUS BUCHLER

DOVER PUBLICATIONS, INC., NEW YORK

7

LOGIC AS SEMIOTIC: THE THEORY OF SIGNS *

1. What is a Sign? Three Divisions of Logic

SEMIOTIC

Logic, in its general sense, is, as I believe I have shown, only another name for *semiotic* (σημειωτική), the quasi-necessary, or formal, doctrine of signs. By describing the doctrine as "quasi-necessary," or formal, I mean that we observe the characters of such signs as we know, and from such an observation, by a process which I will not object to naming Abstraction, we are led to statements, eminently fallible, and therefore in one sense by no means necessary, as to what *must be* the characters of all signs used by a "scientific" intelligence, that is to say, by an intelligence capable of learning by experience. As to that process of abstraction, it is itself a sort of observation. The faculty which I call abstractive observation is one which ordinary people perfectly recognize, but for which the theories of philosophers sometimes hardly leave room. It is a familiar experience to every human being to wish for something quite beyond his present means, and to follow that wish by the question, "Should I wish for that thing just the same, if I had ample means to gratify it?" To answer that question, he searches his heart, and in doing so makes what I term an abstractive observation. He makes in his imagination a sort of skeleton diagram, or outline sketch, of himself, considers what modifications the hypothetical state of things would require to be made in that picture, and then examines it, that is, *observes* what he has imagined, to see whether the same ardent desire is there to be discerned. By such a process, which is at bottom very much like mathematical reasoning, we can reach conclusions as to what *would be* true of signs in all cases, so long as the intelligence using them was scientific. The modes of thought of a God, who should possess an intuitive omniscience superseding reason,

* [The first of the three selections in *1* is from ms. c. 1897 (*CP* 2.227-9), the third from ms. c. 1910 (*CP* 2.231-2). The second selection in *1*, *3b*, the second selection in *3c*, and *3d* are from mss. c. 1902, c. 1895, and c. 1893 (*CP* 2.274-302). *2* and *4* are from ms. c. 1903 (*CP* 2.243-52, 254-65). *3a* is from the article "Sign" in Baldwin's *Dictionary of Philosophy and Psychology* 1902 (*CP* 2.304). The first selection in *3c* is from the article "Index" in Baldwin's (*CP* 2.305, 306).]

are put out of the question. Now the whole process of development among the community of students of those formulations by abstractive observation and reasoning of the truths which *must* hold good of all signs used by a scientific intelligence is an observational science, like any other positive science, notwithstanding its strong contrast to all the special sciences which arises from its aiming to find out what *must be* and not merely what *is* in the actual world.

SIGN

A sign, or *representamen*, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object*. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the *ground* of the representamen. "Idea" is here to be understood in a sort of Platonic sense, very familiar in everyday talk; I mean in that sense in which we say that one man catches another man's idea, in which we say that when a man recalls what he was thinking of at some previous time, he recalls the same idea, and in which when a man continues to think anything, say for a tenth of a second, in so far as the thought continues to agree with itself during that time, that is to have a *like* content, it is the same idea, and is not at each instant of the interval a new idea.

In consequence of every representamen being thus connected with three things, the ground, the object, and the interpretant, the science of semiotic has three branches.[7] The first is called by Duns Scotus *grammatica speculativa*. We may term it *pure grammar*. It has for its task to ascertain what must be true of the representamen used by every scientific intelligence in order that they may embody any *meaning*. The second is logic proper. It is the science of what is quasi-necessarily true of the representamina of any scientific intelligence in order that they may hold good of any *object*, that is, may be true. Or say, logic proper is the formal science of the conditions of the truth of representations. The third, in imitation of Kant's fashion of preserving old associations of words in finding nomenclature for new conceptions, I call *pure rhetoric*. Its task is to ascertain the laws by which in every scientific intelligence one sign gives birth to another, and especially one thought brings forth another.

NB VANNEVAR BUSH

A *Sign*, or *Representamen*, is a First which stands in such a genuine triadic relation to a Second, called its *Object*, as to be capable of

determining a Third, called its *Interpretant*, to assume the same triadic relation to its Object in which it stands itself to the same Object. The triadic relation is *genuine*, that is its three members are bound together by it in a way that does not consist in any complexus of dyadic relations. That is the reason the Interpretant, or Third, cannot stand in a mere dyadic relation to the Object, but must stand in such a relation to it as the Representamen itself does. Nor can the triadic relation in which the Third stands be merely similar to that in which the First stands, for this would make the relation of the Third to the First a degenerate Secondness merely. The Third must indeed stand in such a relation, and thus must be capable of determining a Third of its own; but besides that, it must have a second triadic relation in which the Representamen, or rather the relation thereof to its Object, shall be its own (the Third's) Object, and must be capable of determining a Third to this relation. All this must equally be true of the Third's Thirds and so on endlessly; and this, and more, is involved in the familiar idea of a Sign; and as the term Representamen is here used, nothing more is implied. A *Sign* is a Representamen with a mental Interpretant. Possibly there may be Representamens that are not Signs. Thus, if a sunflower, in turning toward the sun, becomes by that very act fully capable, without further condition, of reproducing a sunflower which turns in precisely corresponding ways toward the sun, and of doing so with the same reproductive power, the sunflower would become a Representamen of the sun. But *thought* is the chief, if not the only, mode of representation.

The Sign can only represent the Object and tell about it. It cannot furnish acquaintance with or recognition of that Object; for that is what is meant in this volume by the Object of a Sign; namely, that with which it presupposes an acquaintance in order to convey some further information concerning it. No doubt there will be readers who will say they cannot comprehend this. They think a Sign need not relate to anything otherwise known, and can make neither head nor tail of the statement that every sign must relate to such an Object. But if there be anything that conveys information and yet has absolutely no relation nor reference to anything with which the person to whom it conveys the information has, when he comprehends that information, the slightest acquaintance, direct or indirect—and a very strange sort of information that would be —the vehicle of that sort of information is not, in this volume, called a Sign.

Two men are standing on the seashore looking out to sea. One of them says to the other, "That vessel there carries no freight at all, but only passengers." Now, if the other, himself, sees no vessel, the first information he derives from the remark has for its Object the part of the sea that he does see, and informs him that a person with sharper eyes than his, or more trained in looking for such things, can see a vessel there; and then, that vessel having been thus introduced to his acquaintance, he is prepared to receive the information about it that it carries passengers exclusively. But the sentence as a whole has, for the person supposed, no other Object than that with which it finds him already acquainted. The Objects—for a Sign may have any number of them—may each be a single known existing thing or thing believed formerly to have existed or expected to exist, or a collection of such things, or a known quality or relation or fact, which single Object may be a collection, or whole of parts, or it may have some other mode of being, such as some act permitted whose being does not prevent its negation from being equally permitted, or something of a general nature desired, required, or invariably found under certain general circumstances.

KINDS OF OBJECTS ✓

2. Three Trichotomies of Signs

Signs are divisible by three trichotomies; first, according as the sign in itself is a mere quality, is an actual existent, or is a general law; secondly, according as the relation of the sign to its object consists in the sign's having some character in itself, or in some existential relation to that object, or in its relation to an interpretant; thirdly, according as its Interpretant represents it as a sign of possibility or as a sign of fact or a sign of reason.

i

According to the first division, a Sign may be termed a *Qualisign*, a *Sinsign*, or a *Legisign*.

A *Qualisign* is a quality which is a Sign. It cannot actually act as a sign until it is embodied; but the embodiment has nothing to do with its character as a sign.

A *Sinsign* (where the syllable *sin* is taken as meaning "being only once," as in *single, simple*, Latin *semel*, etc.) is an actual existent thing or event which is a sign. It can only be so through its qualities; so that it involves a qualisign, or rather, several qualisigns. But these qualisigns are of a peculiar kind and only form a sign through being actually embodied.

A *Legisign* is a law that is a Sign. This law is usually established by men. Every conventional sign is a legisign [but not conversely]. It is not a single object, but a general type which, it has been agreed, shall be significant. Every legisign signifies through an instance of its application, which may be termed a *Replica* of it. Thus, the word "the" will usually occur from fifteen to twenty-five times on a page. It is in all these occurrences one and the same word, the same legisign. Each single instance of it is a Replica. The Replica is a Sinsign. Thus, every Legisign requires Sinsigns. But these are not ordinary Sinsigns, such as are peculiar occurrences that are regarded as significant. Nor would the Replica be significant if it were not for the law which renders it so.

ii

According to the second trichotomy, a Sign may be termed an *Icon*, an *Index*, or a *Symbol*.

An *Icon* is a sign which refers to the Object that it denotes merely by virtue of characters of its own, and which it possesses, just the same, whether any such Object actually exists or not. It is true that unless there really is such an Object, the Icon does not act as a sign; but this has nothing to do with its character as a sign. Anything whatever, be it quality, existent individual, or law, is an Icon of anything, in so far as it is like that thing and used as a sign of it.

An *Index* is a sign which refers to the Object that it denotes by virtue of being really affected by that Object. It cannot, therefore, be a Qualisign, because qualities are whatever they are independently of anything else. In so far as the Index is affected by the Object, it necessarily has some Quality in common with the Object, and it is in respect to these that it refers to the Object. It does, therefore, involve a sort of Icon, although an Icon of a peculiar kind; and it is not the mere resemblance of its Object, even in these respects which makes it a sign, but it is the actual modification of it by the Object.

A *Symbol* is a sign which refers to the Object that it denotes by virtue of a law, usually an association of general ideas, which operates to cause the Symbol to be interpreted as referring to that Object. It is thus itself a general type or law, that is, is a Legisign. As such it acts through a Replica. Not only is it general itself, but the Object to which it refers is of a general nature. Now that which is general has its being in the instances which it will determine. There must, therefore, be existent instances of what the

Symbol denotes, although we must here understand by "existent," existent in the possibly imaginary universe to which the Symbol refers. The Symbol will indirectly, through the association or other law, be affected by those instances; and thus the Symbol will involve a sort of Index, although an Index of a peculiar kind. It will not, however, be by any means true that the slight effect upon the Symbol of those instances accounts for the significant character of the Symbol.

iii

According to the third trichotomy, a Sign may be termed a *Rheme*, a *Dicisign* or *Dicent Sign* (that is, a proposition or quasi-proposition), or an *Argument*.

A *Rheme* is a Sign which, for its Interpretant, is a Sign of qualitative Possibility, that is, is understood as representing such and such a kind of possible Object. Any Rheme, perhaps, will afford some information; but it is not interpreted as doing so.

A *Dicent Sign* is a Sign, which, for its Interpretant, is a Sign of actual existence. It cannot, therefore, be an Icon, which affords no ground for an interpretation of it as referring to actual existence. A Dicisign necessarily involves, as a part of it, a Rheme, to describe the fact which it is interpreted as indicating. But this is a peculiar kind of Rheme; and while it is essential to the Dicisign, it by no means constitutes it.

An *Argument* is a Sign which, for its Interpretant, is a Sign of law. Or we may say that a Rheme is a sign which is understood to represent its object in its characters merely; that a Dicisign is a sign which is understood to represent its object in respect to actual existence; and that an Argument is a Sign which is understood to represent its Object in its character as Sign. Since these definitions touch upon points at this time much in dispute, a word may be added in defence of them. A question often put is: What is the essence of a Judgment? A judgment is the mental act by which the judger seeks to impress upon himself the truth of a proposition. It is much the same as an act of asserting the proposition, or going before a notary and assuming formal responsibility for its truth, except that those acts are intended to affect others, while the judgment is only intended to affect oneself. However, the logician, as such, cares not what the psychological nature of the act of judging may be. The question for him is: What is the nature of the sort of sign of which a principal variety is called a proposition, which is the matter upon which the act of judging is exercised? The pro-

position need not be asserted or judged. It may be contemplated as a sign capable of being asserted or denied. This sign itself retains its full meaning whether it be actually asserted or not. The peculiarity of it, therefore, lies in its mode of meaning; and to say this is to say that its peculiarity lies in its relation to its interpretant. The proposition professes to be really affected by the actual existent or real law to which it refers. The argument makes the same pretension, but that is not the principal pretension of the argument. The rheme makes no such pretension.

3. Icon, Index, and Symbol

a. Synopsis

A sign is either an *icon*, an *index*, or a *symbol*. An *icon* is a sign which would possess the character which renders it significant, even though its object had no existence; such as a lead-pencil streak as representing a geometrical line. An *index* is a sign which would, at once, lose the character which makes it a sign if its object were removed, but would not lose that character if there were no interpretant. Such, for instance, is a piece of mould with a bullet-hole in it as sign of a shot; for without the shot there would have been no hole; but there is a hole there, whether anybody has the sense to attribute it to a shot or not. A *symbol* is a sign which would lose the character which renders it a sign if there were no interpretant. Such is any utterance of speech which signifies what it does only by virtue of its being understood to have that signification.

b. Icon

. . . While no Representamen actually functions as such until it actually determines an Interpretant, yet it becomes a Representamen as soon as it is fully capable of doing this; and its Representative Quality is not necessarily dependent upon its ever actually determining an Interpretant, nor even upon its actually having an Object.

An *Icon* is a Representamen whose Representative Quality is a Firstness of it as a First. That is, a quality that it has *qua* thing renders it fit to be a representamen. Thus, anything is fit to be a *Substitute* for anything that it is like. (The conception of "substitute" involves that of a purpose, and thus of genuine thirdness.) Whether there are other kinds of substitutes or not we shall see. A Representamen by Firstness alone can only have a similar

Object. Thus, a Sign by Contrast denotes its object only by virtue of a contrast, or Secondness, between two qualities. A sign by Firstness is an image of its object and, more strictly speaking, can only be an *idea*. For it must produce an Interpretant idea; and an external object excites an idea by a reaction upon the brain. But most strictly speaking, even an idea, except in the sense of a possibility, or Firstness, cannot be an Icon. A possibility alone is an Icon purely by virtue of its quality; and its object can only be a Firstness. But a sign may be *iconic*, that is, may represent its object mainly by its similarity, no matter what its mode of being. If a substantive be wanted, an iconic representamen may be termed a *hypoicon*. Any material image, as a painting, is largely conventional in its mode of representation; but in itself, without legend or label it may be called a *hypoicon*.

Hypoicons may be roughly divided according to the mode of Firstness of which they partake. Those which partake of simple qualities, or First Firstnesses, are *images*; those which represent the relations, mainly dyadic, or so regarded, of the parts of one thing by analogous relations in their own parts, are *diagrams*; those which represent the representative character of a representamen by representing a parallelism in something else, are *metaphors*.

The only way of directly communicating an idea is by means of an icon; and every indirect method of communicating an idea must depend for its establishment upon the use of an icon. Hence, every assertion must contain an icon or set of icons, or else must contain signs whose meaning is only explicable by icons. The idea which the set of icons (or the equivalent of a set of icons) contained in an assertion signifies may be termed the *predicate* of the assertion.

Turning now to the rhetorical evidence, it is a familiar fact that there are such representations as icons. Every picture (however conventional its method) is essentially a representation of that kind. So is every diagram, even although there be no sensuous resemblance between it and its object, but only an analogy between the relations of the parts of each. Particularly deserving of notice are icons in which the likeness is aided by conventional rules. Thus, an algebraic formula is an icon, rendered such by the rules of commutation, association, and distribution of the symbols. It may seem at first glance that it is an arbitrary classification to call an algebraic expression an icon; that it might as well, or better, be regarded as a compound conventional sign. But it is not so. For a great distinguishing property of the icon is that by the direct observation of it other truths concerning its object can be dis-

covered than those which suffice to determine its construction. Thus, by means of two photographs a map can be drawn, etc. Given a conventional or other general sign of an object, to deduce any other truth than that which it explicitly signifies, it is necessary, in all cases, to replace that sign by an icon. This capacity of revealing unexpected truth is precisely that wherein the utility of algebraical formulae consists, so that the iconic character is the prevailing one.

That icons of the algebraic kind, though usually very simple ones, exist in all ordinary grammatical propositions is one of the philosophic truths that the Boolean logic brings to light. In all primitive writing, such as the Egyptian hieroglyphics, there are icons of a non-logical kind, the ideographs. In the earliest form of speech, there probably was a large element of mimicry. But in all languages known, such representations have been replaced by conventional auditory signs. These, however, are such that they can only be explained by icons. But in the syntax of every language there are logical icons of the kind that are aided by conventional rules. . . .

Photographs, especially instantaneous photographs, are very instructive, because we know that they are in certain respects exactly like the objects they represent. But this resemblance is due to the photographs having been produced under such circumstances that they were physically forced to correspond point by point to nature. In that aspect, then, they belong to the second class of signs, those by physical connection. The case is different if I surmise that zebras are likely to be obstinate, or otherwise disagreeable animals, because they seem to have a general resemblance to donkeys, and donkeys are self-willed. Here the donkey serves precisely as a probable likeness of the zebra. It is true we suppose that resemblance has a physical cause in heredity; but then, this hereditary affinity is itself only an inference from the likeness between the two animals, and we have not (as in the case of the photograph) any independent knowledge of the circumstances of the production of the two species. Another example of the use of a likeness is the design an artist draws of a statue, pictorial composition, architectural elevation, or piece of decoration, by the contemplation of which he can ascertain whether what he proposes will be beautiful and satisfactory. The question asked is thus answered almost with certainty because it relates to how the artist will himself be affected. The reasoning of mathematicians will be found to turn chiefly upon the use of likenesses, which are the very hinges of the gates of their science. The utility of likenesses to

mathematicians consists in their suggesting in a very precise way, new aspects of supposed states of things. . . .

Many diagrams resemble their objects not at all in looks; it is only in respect to the relations of their parts that their likeness consists. Thus, we may show the relation between the different kinds of signs by a brace, thus:

$$\text{Signs: } \begin{cases} \text{Icons,} \\ \text{Indices,} \\ \text{Symbols.} \end{cases}$$

This is an icon. But the only respect in which it resembles its object is that the brace shows the classes of *icons*, *indices*, and *symbols* to be related to one another and to the general class of signs, as they really are, in a general way. When, in algebra, we write equations under one another in a regular array, especially when we put resembling letters for corresponding coefficients, the array is an icon. Here is an example:

$$a_1x+b_1y=n_1,$$
$$a_2x+b_2y=n_2.$$

This is an icon,[8] in that it makes quantities look alike which are in analogous relations to the problem. In fact, every algebraical equation is an icon, in so far as it *exhibits*, by means of the algebraical signs (which are not themselves icons), the relations of the quantities concerned.

It may be questioned whether all icons are likenesses or not. For example, if a drunken man is exhibited in order to show, by contrast, the excellence of temperance, this is certainly an icon, but whether it is a likeness or not may be doubted. The question seems somewhat trivial.

c. Index

[An index is] a sign, or representation, which refers to its object not so much because of any similarity or analogy with it, nor because it is associated with general characters which that object happens to possess, as because it is in dynamical (including spatial) connection both with the individual object, on the one hand, and with the senses or memory of the person for whom it serves as a sign, on the other hand. . . . While demonstrative and personal pronouns are, as ordinarily used, "genuine indices," relative pronouns are "degenerate indices"; for though they may, accidentally and indirectly, refer to existing things, they directly refer,

and need only refer, to the images in the mind which previous words have created.

Indices may be distinguished from other signs, or representations, by three characteristic marks: first, that they have no significant resemblance to their objects; second, that they refer to individuals, single units, single collections of units, or single continua; third, that they direct the attention to their objects by blind compulsion. But it would be difficult, if not impossible, to instance an absolutely pure index, or to find any sign absolutely devoid of the indexical quality. Psychologically, the action of indices depends upon association by contiguity, and not upon association by resemblance or upon intellectual operations.

An *Index* or *Seme* ($\sigma\hat{\eta}\mu\alpha$) is a Representamen whose Representative character consists in its being an individual second. If the Secondness is an existential relation, the Index is *genuine*. If the Secondness is a reference, the Index is *degenerate*. A genuine Index and its Object must be existent individuals (whether things or facts), and its immediate Interpretant must be of the same character. But since every individual must have characters, it follows that a genuine Index may contain a Firstness, and so an Icon as a constituent part of it. Any individual is a degenerate Index of its own characters.

Subindices or *Hyposemes* are signs which are rendered such principally by an actual connection with their objects. Thus a proper name, personal demonstrative, or relative pronoun or the letter attached to a diagram, denotes what it does owing to a real connection with its object, but none of these is an Index, since it is not an individual.

Let us examine some examples of indices. I see a man with a rolling gait. This is a probable indication that he is a sailor. I see a bowlegged man in corduroys, gaiters, and a jacket. These are probable indications that he is a jockey or something of the sort. A sundial or a clock *indicates* the time of day. Geometricians mark letters against the different parts of their diagrams and then use these letters to indicate those parts. Letters are similarly used by lawyers and others. Thus, we may say: If A and B are married to one another and C is their child while D is brother of A, then D is uncle of C. Here A, B, C, and D fulfill the office of relative pronouns, but are more convenient since they require no special collocation of words. A rap on the door is an index. Anything which focusses the attention is an index. Anything which startles us is an index,

in so far as it marks the junction between two portions of experience. Thus a tremendous thunderbolt indicates that *something* considerable happened, though we may not know precisely what the event was. But it may be expected to connect itself with some other experience.

. . . A low barometer with a moist air is an index of rain; that is we suppose that the forces of nature establish a probable connection between the low barometer with moist air and coming rain. A weathercock is an index of the direction of the wind; because in the first place it really takes the self-same direction as the wind, so that there is a real connection between them, and in the second place we are so constituted that when we see a weathercock pointing in a certain direction it draws our attention to that direction, and when we see the weathercock veering with the wind, we are forced by the law of mind to think that direction is connected with the wind. The pole star is an index, or pointing finger, to show us which way is north. A spirit-level, or a plumb bob, is an index of the vertical direction. A yard-stick might seem, at first sight, to be an icon of a yard; and so it would be, if it were merely intended to show a yard as near as it can be seen and estimated to be a yard. But the very purpose of a yard-stick is to show a yard nearer than it can be estimated by its appearance. This it does in consequence of an accurate mechanical comparison made with the bar in London called the yard. Thus it is a real connection which gives the yard-stick its value as a representamen; and thus it is an *index*, not a mere *icon*.

When a driver to attract the attention of a foot passenger and cause him to save himself, calls out "Hi!" so far as this is a significant word, it is, as will be seen below, something more than an index; but so far as it is simply intended to act upon the hearer's nervous system and to rouse him to get out of the way, it is an index, because it is meant to put him in real connection with the object, which is his situation relative to the approaching horse. Suppose two men meet upon a country road and one of them says to the other, "The chimney of that house is on fire." The other looks about him and descries a house with green blinds and a verandah having a smoking chimney. He walks on a few miles and meets a second traveller. Like a Simple Simon he says, "The chimney of that house is on fire." "What house?" asks the other. "Oh, a house with green blinds and a verandah," replies the simpleton. "Where is the house?" asks the stranger. He desires some *index* which shall connect his apprehension with the house meant.

Words alone cannot do this. The demonstrative pronouns, "this" and "that," are indices. For they call upon the hearer to use his powers of observation, and so establish a real connection between his mind and the object; and if the demonstrative pronoun does that—without which its meaning is not understood—it goes to establish such a connection; and so is an index. The relative pronouns, *who* and *which*, demand observational activity in much the same way, only with them the observation has to be directed to the words that have gone before. Lawyers use A, B, C, practically as very effective relative pronouns. To show how effective they are, we may note that Messrs. Allen and Greenough, in their admirable (though in the edition of 1877 [?], too small) Latin Grammar, declare that no conceivable syntax could wholly remove the ambiguity of the following sentence, "A replied to B that he thought C (his brother) more unjust to himself than to his own friend." Now, any lawyer would state that with perfect clearness, by using A, B, C, as relatives, thus:

A replied to B that he $\left\{\begin{matrix}\text{A}\\\text{B}\end{matrix}\right\}$, thought C (his $\left\{\begin{matrix}\text{A's}\\\text{B's}\end{matrix}\right\}$, brother) more unjust to himself, $\left\{\begin{matrix}\text{A}\\\text{B}\\\text{C}\end{matrix}\right\}$ than to his $\left\{\begin{matrix}\text{A's}\\\text{B's}\\\text{C's}\end{matrix}\right\}$ own friend. The terminations which in any inflected language are attached to words "governed" by other words, and which serve to show which the governing word is, by repeating what is elsewhere expressed in the same form, are likewise *indices* of the same relative pronoun character. Any bit of Latin poetry illustrates this, such as the twelve-line sentence beginning, "*Jam satis terris.*" Both in these terminations and in the A, B, C, a likeness is relied upon to carry the attention to the right object. But this does not make them icons, in any important way; for it is of no consequence how the letters A, B, C, are shaped or what the terminations are. It is not merely that one occurrence of an A is like a previous occurrence that is the important circumstance, but that *there is an understanding that like letters shall stand for the same thing*, and this acts as a force carrying the attention from one occurrence of A to the previous one. A possessive pronoun is two ways an index: first it indicates the possessor, and, second, it has a modification which syntactically carries the attention to the word denoting the thing possessed.

Some indices are more or less detailed directions for what the hearer is to do in order to place himself in direct experiential or other connection with the thing meant. Thus, the Coast Survey issues

"Notices to Mariners," giving the latitude and longitude, four or five bearings of prominent objects, etc., and saying *there* is a rock, or shoal, or buoy, or lightship. Although there will be other elements in such directions, yet in the main they are indices.

Along with such indexical directions of what to do to find the object meant, ought to be classed those pronouns which should be entitled *selective* pronouns [or quantifiers] because they inform the hearer how he is to pick out one of the objects intended, but which grammarians call by the very indefinite designation of *indefinite* pronouns. Two varieties of these are particularly important in logic, the *universal selectives*, such as *quivis, quilibet, quisquam, ullus, nullus, nemo, quisque, uterque*, and in English, *any, every, all, no, none, whatever, whoever, everybody, anybody, nobody*. These mean that the hearer is at liberty to select any instance he likes within limits expressed or understood, and the assertion is intended to apply to that one. The other logically important variety consists of the *particular selectives, quis, quispiam, nescio quis, aliquis, quidam*, and in English, *some, something, somebody, a, a certain, some or other, a suitable, one.*

Allied to the above pronouns are such expressions as *all but one, one or two, a few, nearly all, every other one*, etc. Along with pronouns are to be classed adverbs of place and time, etc.

Not very unlike these are, *the first, the last, the seventh, two-thirds of, thousands of*, etc.

Other indexical words are prepositions, and prepositional phrases, such as, "on the right (or left) of." Right and left cannot be distinguished by any general description. Other prepositions signify relations which may, perhaps, be described; but when they refer, as they do oftener than would be supposed, to a situation relative to the observed, or assumed to be experientially known, place and attitude of the speaker relatively to that of the hearer, then the indexical element is the dominant element.

Icons and indices assert nothing. If an icon could be interpreted by a sentence, that sentence must be in a "potential mood," that is, it would merely say, "Suppose a figure has three sides," etc. Were an index so interpreted, the mood must be imperative, or exclamatory, as "See there!" or "Look out!" But the kind of signs which we are now coming to consider are, by nature, in the "indicative," or, as it should be called, the *declarative* mood. Of course, they can go to the expression of any other mood, since we may declare assertions to be doubtful, or mere interrogations, or imperatively requisite.

d. Symbol

A Symbol is a Representamen whose Representative character consists precisely in its being a rule that will determine its Interpretant. All words, sentences, books, and other conventional signs are Symbols. We speak of writing or pronouncing the word "man"; but it is only a *replica*, or embodiment of the word, that is pronounced or written. The word itself has no existence although it has a real being, *consisting in* the fact that existents *will* conform to it. It is a general mode of succession of three sounds or representamens of sounds, which becomes a sign only in the fact that a habit, or acquired law, will cause replicas of it to be interpreted as meaning a man or men. The word and its meaning are both general rules; but the word alone of the two prescribes the qualities of its replicas in themselves. Otherwise the "word" and its "meaning" do not differ, unless some special sense be attached to "meaning."

A Symbol is a law, or regularity of the indefinite future. Its Interpretant must be of the same description; and so must be also the complete immediate Object, or meaning. But a law necessarily governs, or "is embodied in" individuals, and prescribes some of their qualities. Consequently, a constituent of a Symbol may be an Index, and a constituent may be an Icon. A man walking with a child points his arm up into the air and says, "There is a balloon." The pointing arm is an essential part of the symbol without which the latter would convey no information. But if the child asks, "What is a balloon," and the man replies, "It is something like a great big soap bubble," he makes the image a part of the symbol. Thus, while the complete object of a symbol, that is to say, its meaning, is of the nature of a law, it must *denote* an individual, and must *signify* a character. A *genuine* symbol is a symbol that has a general meaning. There are two kinds of degenerate symbols, the *Singular Symbol* whose Object is an existent individual, and which signifies only such characters as that individual may realize; and the *Abstract Symbol*, whose only Object is a character.

Although the immediate Interpretant of an Index must be an Index, yet since its Object may be the Object of an Individual [Singular] Symbol, the Index may have such a Symbol for its indirect Interpretant. Even a genuine Symbol may be an imperfect Interpretant of it. So an *icon* may have a degenerate Index, or an Abstract Symbol, for an indirect Interpretant, and a genuine Index or Symbol for an imperfect Interpretant.

A *Symbol* is a sign naturally fit to declare that the set of objects

which is denoted by whatever set of indices may be in certain ways attached to it is represented by an icon associated with it. To show what this complicated definition means, let us take as an example of a symbol the word "loveth." Associated with this word is an idea, which is the mental icon of one person loving another. Now we are to understand that "loveth" occurs in a sentence; for what it may mean by itself, if it means anything, is not the question. Let the sentence, then, be "Ezekiel loveth Huldah." Ezekiel and Huldah must, then, be or contain indices; for without indices it is impossible to designate what one is talking about. Any mere description would leave it uncertain whether they were not mere characters in a ballad; but whether they be so or not, indices can designate them. Now the effect of the word "loveth" is that the pair of objects denoted by the pair of indices Ezekiel and Huldah is represented by the icon, or the image we have in our minds of a lover and his beloved.

The same thing is equally true of every verb in the declarative mood; and indeed of every verb, for the other moods are merely declarations of a fact somewhat different from that expressed by the declarative mood. As for a noun, considering the meaning which it has in the sentence, and not as standing by itself, it is most conveniently regarded as a portion of a symbol. Thus the sentence, "every man loves a woman" is equivalent to "whatever is a man loves something that is a woman." Here "whatever" is a universal selective index, "is a man" is a symbol, "loves" is a symbol, "something that" is a particular selective index, and "is a woman" is a symbol. . . .

The word *Symbol* has so many meanings that it would be an injury to the language to add a new one. I do not think that the signification I attach to it, that of a conventional sign, or one depending upon habit (acquired or inborn), is so much a new meaning as a return to the original meaning. Etymologically, it should mean a thing thrown together, just as ἔμβολον (embolum) is a thing thrown into something, a bolt, and παράβολον (parabolum) is a thing thrown besides, collateral security, and ὑπόβολον (hypobolum) is a thing thrown underneath, an antenuptial gift. It is usually said that in the word *symbol* the throwing together is to be understood in the sense of "to conjecture"; but were that the case, we ought to find that *sometimes* at least it meant a conjecture, a meaning for which literature may be searched in vain. But the Greeks used "throw together" (συμβάλλειν) very frequently to signify the making of a contract or convention. Now, we do find symbol (σύμβολον) early

and often used to mean a convention or contract. Aristotle calls a noun a "symbol," that is, a conventional sign. In Greek, watch-fire is a "symbol," that is, a signal agreed upon; a standard or ensign is a "symbol," a watchword is a "symbol," a badge is a "symbol"; a church creed is called a "symbol," because it serves as a badge or shibboleth; a theatre ticket is called a "symbol"; any ticket or check entitling one to receive anything is a "symbol." Moreover, any expression of sentiment was called a "symbol." Such were the principal meanings of the word in the original language. The reader will judge whether they suffice to establish my claim that I am not seriously wrenching the word in employing it as I propose to do.

Any ordinary word, as "give," "bird," "marriage," is an example of a symbol. It is *applicable to whatever may be found to realize the idea connected with the word*; it does not, in itself, identify those things. It does not show us a bird, nor enact before our eyes a giving or a marriage, but supposes that we are able to imagine those things, and have associated the word with them.

A regular progression of one, two, three may be remarked in the three orders of signs, Icon, Index, Symbol. The Icon has no dynamical connection with the object it represents; it simply happens that its qualities resemble those of that object, and excite analogous sensations in the mind for which it is a likeness. But it really stands unconnected with them. The index is physically connected with its object; they make an organic pair, but the interpreting mind has nothing to do with this connection, except remarking it, after it is established. The symbol is connected with its object by virtue of the idea of the symbol-using mind, without which no such connection would exist.

Every physical force reacts between a pair of particles, either of which may serve as an index of the other. On the other hand, we shall find that every intellectual operation involves a triad of symbols.

A symbol, as we have seen, cannot indicate any particular thing; it denotes a kind of thing. Not only that, but it is itself a kind and not a single thing. You can write down the word "star," but that does not make you the creator of the word, nor if you erase it have you destroyed the word. The word lives in the minds of those who use it. Even if they are all asleep, it exists in their memory. So we may admit, if there be reason to do so, that generals are mere words without at all saying, as Ockham supposed, that they are really individuals.

Symbols grow. They come into being by development out of other signs, particularly from icons, or from mixed signs partaking of the nature of icons and symbols. We think only in signs. These mental signs are of mixed nature; the symbol-parts of them are called concepts. If a man makes a new symbol, it is by thoughts involving concepts. So it is only out of symbols that a new symbol can grow. *Omne symbolum de symbolo.* A symbol, once in being, spreads among the peoples. In use and in experience, its meaning grows. Such words as *force, law, wealth, marriage,* bear for us very different meanings from those they bore to our barbarous ancestors. The symbol may, with Emerson's sphynx, say to man,

> Of thine eye I am eyebeam.

4. Ten Classes of Signs

The three trichotomies of Signs result together in dividing Signs into TEN CLASSES OF SIGNS, of which numerous subdivisions have to be considered. The ten classes are as follows:

First: A Qualisign [*e.g.*, a feeling of "red"] is any quality in so far as it is a sign. Since a quality is whatever it is positively in itself, a quality can only denote an object by virtue of some common ingredient or similarity; so that a Qualisign is necessarily an Icon. Further, since a quality is a mere logical possibility, it can only be interpreted as a sign of essence, that is, as a Rheme.

Second: An Iconic Sinsign [*e.g.*, an individual diagram] is any object of experience in so far as some quality of it makes it determine the idea of an object. Being an Icon, and thus a sign by likeness purely, of whatever it may be like, it can only be interpreted as a sign of essence, or Rheme. It will embody a Qualisign.

Third: A Rhematic Indexical Sinsign [*e.g.*, a spontaneous cry] is any object of direct experience so far as it directs attention to an Object by which its presence is caused. It necessarily involves an Iconic Sinsign of a peculiar kind, yet is quite different since it brings the attention of the interpreter to the very Object denoted.

Fourth: A Dicent Sinsign [*e.g.*, a weathercock] is any object of direct experience, in so far as it is a sign, and, as such, affords information concerning its Object. This it can only do by being really affected by its Object; so that it is necessarily an Index. The only information it can afford is of actual fact. Such a Sign must involve an Iconic Sinsign to embody the information and a Rhematic Indexical Sinsign to indicate the Object to which the

information refers. But the mode of combination, or *Syntax*, of these two must also be significant.

Fifth: An Iconic Legisign [*e.g.*, a diagram, apart from its factual individuality] is any general law or type, in so far as it requires each instance of it to embody a definite quality which renders it fit to call up in the mind the idea of a like object. Being an Icon, it must be a Rheme. Being a Legisign, its mode of being is that of governing single Replicas, each of which will be an Iconic Sinsign of a peculiar kind.

Sixth: A Rhematic Indexical Legisign [*e.g.*, a demonstrative pronoun] is any general type or law, however established, which requires each instance of it to be really affected by its Object in such a manner as merely to draw attention to that Object. Each Replica of it will be a Rhematic Indexical Sinsign of a peculiar kind. The Interpretant of a Rhematic Indexical Legisign represents it as an Iconic Legisign; and so it is, in a measure—but in a very small measure.

Seventh: A Dicent Indexical Legisign [*e.g.*, a street cry] is any general type or law, however established, which requires each instance of it to be really affected by its Object in such a manner as to furnish definite information concerning that Object. It must involve an Iconic Legisign to signify the information and a Rhematic Indexical Legisign to denote the subject of that information. Each Replica of it will be a Dicent Sinsign of a peculiar kind.

Eighth: A Rhematic Symbol or Symbolic Rheme [*e.g.*, a common noun] is a sign connected with its Object by an association of general ideas in such a way that its Replica calls up an image in the mind, which image, owing to certain habits or dispositions of that mind, tends to produce a general concept, and the Replica is interpreted as a Sign of an Object that is an instance of that concept. Thus, the Rhematic Symbol either is, or is very like, what the logicians call a General Term. The Rhematic Symbol, like any Symbol, is necessarily itself of the nature of a general type, and is thus a Legisign. Its Replica, however, is a Rhematic Indexical Sinsign of a peculiar kind, in that the image it suggests to the mind acts upon a Symbol already in that mind to give rise to a General Concept. In this it differs from other Rhematic Indexical Sinsigns, including those which are Replicas of Rhematic Indexical Legisigns. Thus, the demonstrative pronoun "that" is a Legisign, being a general type; but it is not a Symbol, since it does not signify a general concept. Its Replica draws attention to a single Object, and is a Rhematic Indexical Sinsign. A Replica of the

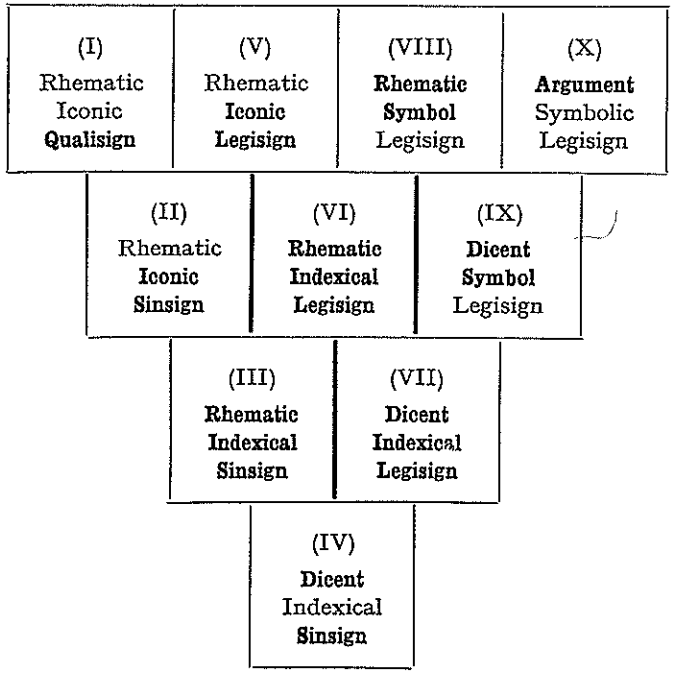
word "camel" is likewise a Rhematic Indexical Sinsign, being really affected, through the knowledge of camels, common to the speaker and auditor, by the real camel it denotes, even if this one is not individually known to the auditor; and it is through such real connection that the word "camel" calls up the idea of a camel. The same thing is true of the word "phoenix." For although no phoenix really exists, real descriptions of the phoenix are well known to the speaker and his auditor; and thus the word is really affected by the Object denoted. But not only are the Replicas of Rhematic Symbols very different from ordinary Rhematic Indexical Sinsigns, but so likewise are Replicas of Rhematic Indexical Legisigns. For the thing denoted by "that" has not affected the replica of the word in any such direct and simple manner as that in which, for example, the ring of a telephone-bell is affected by the person at the other end who wants to make a communication. The Interpretant of the Rhematic Symbol often represents it as a Rhematic Indexical Legisign; at other times as an Iconic Legisign; and it does in a small measure partake of the nature of both.

Ninth: A Dicent Symbol, or ordinary Proposition, is a sign connected with its object by an association of general ideas, and acting like a Rhematic Symbol, except that its intended interpretant represents the Dicent Symbol as being, in respect to what it signifies, really affected by its Object, so that the existence or law which it calls to mind must be actually connected with the indicated Object. Thus, the intended Interpretant looks upon the Dicent Symbol as a Dicent Indexical Legisign; and if it be true, it does partake of this nature, although this does not represent its whole nature. Like the Rhematic Symbol, it is necessarily a Legisign. Like the Dicent Sinsign it is composite inasmuch as it necessarily involves a Rhematic Symbol (and thus is for its Interpretant an Iconic Legisign) to express its information and a Rhematic Indexical Legisign to indicate the subject of that information. But its Syntax of these is significant. The Replica of the Dicent Symbol is a Dicent Sinsign of a peculiar kind. This is easily seen to be true when the information the Dicent Symbol conveys is of actual fact. When that information is of a real law, it is not true in the same fullness. For a Dicent Sinsign cannot convey information of law. It is, therefore, true of the Replica of such a Dicent Symbol only in so far as the law has its being in instances.

Tenth: An Argument is a sign whose interpretant represents its object as being an ulterior sign through a law, namely, the law that the passage from all such premisses to such conclusions tends to

the truth. Manifestly, then, its object must be general; that is, the Argument must be a Symbol. As a Symbol it must, further, be a Legisign. Its Replica is a Dicent Sinsign.

The affinities of the ten classes are exhibited by arranging their designations in the triangular table here shown, which has heavy boundaries between adjacent squares that are appropriated to classes alike in only one respect. All other adjacent squares pertain to classes alike in two respects. Squares not adjacent pertain to classes alike in one respect only, except that each of the three squares of the vertices of the triangle pertains to a class differing in all three respects from the classes to which the squares along the opposite side of the triangle are appropriated. The lightly printed designations are superfluous.



In the course of the above descriptions of the classes, certain subdivisions of some of them have been directly or indirectly referred to. Namely, beside the normal varieties of Sinsigns, Indices, and Dicisigns, there are others which are Replicas of Legisigns, Symbols, and Arguments, respectively. Beside the normal

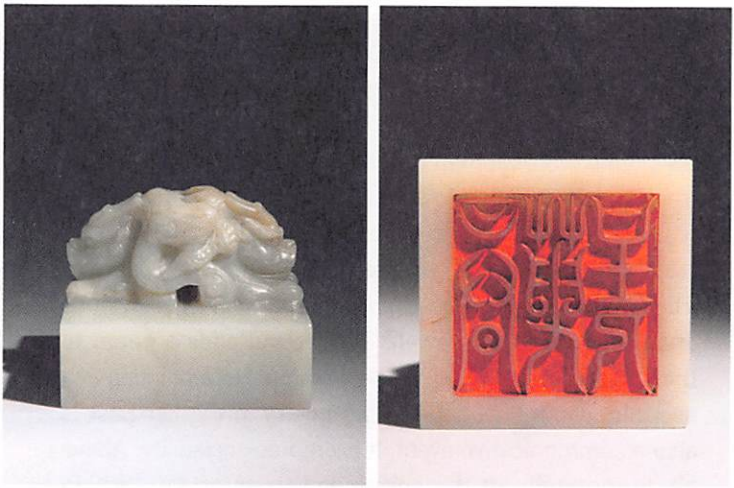
varieties of Qualisigns, Icons, and Rhemes, there are two series of others; to wit, those which are directly involved in Sinsigns, Indices, and Dicisigns, respectively, and also those which are indirectly involved in Legisigns, Symbols, and Arguments, respectively. Thus, the ordinary Dicent Sinsign is exemplified by a weathercock and its veering and by a photograph. The fact that the latter is known to be the effect of the radiations from the object renders it an index and highly informative. A second variety is a Replica of a Dicent Indexical Legisign. Thus any given street cry, since its tone and theme identifies the individual, is not a symbol, but an Indexical Legisign; and any individual instance of it is a Replica of it which is a Dicent Sinsign. A third variety is a Replica of a Proposition. A fourth variety is a Replica of an Argument. Beside the normal variety of the Dicent Indexical Legisign, of which a street cry is an example, there is a second variety, which is that sort of proposition which has the name of a well-known individual as its predicate; as if one is asked, "Whose statue is this?" the answer may be, "It is Farragut." The meaning of this answer is a Dicent Indexical Legisign. A third variety may be a premiss of an argument. A Dicent Symbol, or ordinary proposition, in so far as it is a premiss of an Argument, takes on a new force, and becomes a second variety of the Dicent Symbol. It would not be worth while to go through all the varieties; but it may be well to consider the varieties of one class more. We may take the Rhematic Indexical Legisign. *The* shout of "Hullo!" is an example of the ordinary variety—meaning, not an individual shout, but this shout "Hullo!" in general—this type of shout. A second variety is a constituent of a Dicent Indexical Legisign; as the word "that" in the reply, "that is Farragut." A third variety is a particular application of a Rhematic Symbol; as the exclamation "Hark!" A fourth and fifth variety are in the peculiar force a general word may have in a proposition or argument. It is not impossible that some varieties are here overlooked. It is a nice problem to say to what class a given sign belongs; since all the circumstances of the case have to be considered. But it is seldom requisite to be very accurate; for if one does not locate the sign precisely, one will easily come near enough to its character for any ordinary purpose of logic.

XIAN ZHAN: SEALS OF LEISURE

閑章

The seal stamps on old Chinese paintings are fundamentally different from the signatures used in European painting. Primarily they do not express the authorship that might have authenticated the picture, thereby making it unassailable. Instead, most seal stamps come from the connoisseurs or collectors who inscribe themselves into the picture not only through their seals but also through their commentaries. Here art is a communicative, interactive practice that constantly changes even the artwork's appearance. Subsequent viewers of the picture take part in its creation. The more famous a picture, the more its fame makes it subject to alterations. In addition, the history of a collection is art-historically important in that it changes the work both physically and aesthetically.[1]

In Chinese painting, seal stamps themselves form part of the picture's composition. Thus they are not a paratext but belong to the text itself. From the start Chinese paintings are designed to facilitate later inscriptions. With areas of the picture left empty as communicative spaces, they directly invite viewers to inscribe themselves. Thus with his seal the Chinese painter does not establish his *presence* as a creative subjectivity. Rather, he uses it to open a field of dialogue by merely marking a *trace* that serves to take it forward. As well as name and location seals (名章), there are also so-called *seals of leisure* (閑章). They contain beautiful aphorisms with poetic or moral content. The art-loving Emperor Qianlong (乾隆, 1711–99) is said to have owned some 1,000 seals, from a small seal of just 4 mm with the inscription "ancient fragrance" (古香), to a seal of



Qianlong's jade seal:
"The Ruler Who Believes
in Heaven."

over 20 cm that bears an entire poem in praise of virtue. One seal, which he is said to have used after his enthronement as emperor, reads: "It is hard to be a ruler" (爲君難).

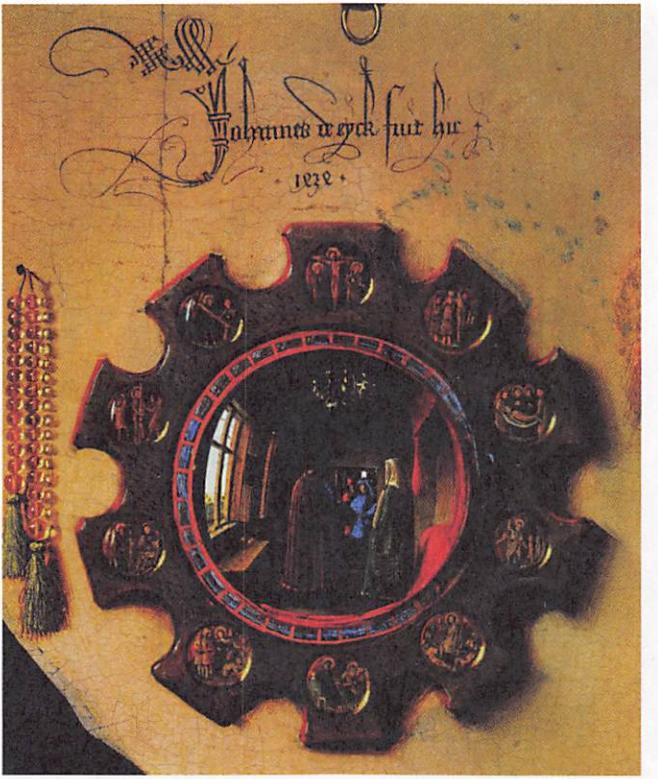
According to a charming custom among Chinese officials, who were all men of letters,[2] a friend, who is being transferred to another location, is taken to a beautiful country spot to celebrate his departure. A picture by Wang Fu (1362–1416), titled *Farewell Meeting at Feng-ch'eng*, depicts a beautiful mountain landscape with a pavilion where friends are celebrating his farewell. Each friend adds a poem to the landscape picture with a seal stamp. Here painting is a sociable, communal act. Writing poetry in the Far East is also a communicative event. It promotes sociability. Above all, it serves to lift the spirits and to entertain. Writing poetry is thus not based on the suffering of a solitary, poetic soul.

Wang Fu, *Farewell Meeting at Feng-ch'eng.*

The seal stamps on Chinese paintings do not actually *finalize* anything. Rather, they *open up* a communicative space. They lend the picture no authorial, authoritative presence. In this they differ significantly from the signatures used in European painting.[3] As seals of *finito*, the latter finalize the *work*; indeed, it is as if they seal it up and prohibit interventions. In contrast to Chinese seal stamps, which are inclusive and communicative, they have an exclusive, executive effect.

Van Eyck's picture, *The Arnolfini Portrait*, embodies the *image of presence*, which is diametrically opposed to the Chinese *image of absence*. The signature "Johannes de Eyck fuit hic" ("Jan van Eyck was here" or "Jan van Eyck painted this") in the middle of the picture locates the presence of the painter in the center of the image. The consciously staged simultaneity of authorship and bearing witness consolidates and concentrates his presence.[4] The signature lends the picture the character of an inalterable, definitive document. The date 1434 below the signature *fixes* the picture in time. Thus any alteration would amount to a falsification of the *truth*. In addition, the demonstrative pronoun *hic* creates a referential structure that specifically emphasizes authorship in a way that the name alone would not be capable of. It expressly points out that the named person is also the creator of the artwork.

Jan van Eyck, *The Arnolfini Portrait*.



The "Eye of God."

Below the signature is a convex mirror. Its radius of reflection extends beyond the frame of the main picture. This creates the impression that the mirror is reflecting reality, whose components are mimetically depicted by the picture. Thus the painting offers itself as a mirror on the world. In the mirror we can see two more people who are present at the betrothal scene. The signature "Johannes de Eyck fuit hic" immediately above the mirror suggests that van Eyck is one of these spectators. Thus the painter is present not only as a signature but also as an image within the image. The mirror is the locus of self-reflection for the painter and for painting. These multiple authorial inscriptions emphasize him specifically as the creator of the picture.

Moreover, because of its circular shape and lateral light reflections, the convex mirror looks like an eye. The medallions that frame the mirror and depict scenes from the Passion, and the cross shape of the window reflected in it, create a clear link between the mirror and the eye of Christ. The divine gaze that coincides in the center of the picture with the scriptural and figurative presence of the painter consolidates the structure of subjectivity. The Passion of Christ is commonly considered to be a mirror of the soul. In this way the painting is *inspired* (*be-seelt*) by the juxtaposition of representations of the Passion and the real mirror. The picture's true setting is thus the *soul.*

Chinese images of absence are, by contrast, *without soul.* Neither authorship nor bearing witness attaches them to identity. In addition, as a result of their aperspectivity and asubjectivity, they are *gaze-less.* In *Berlin Childhood around*

1900, Walter Benjamin recounts an anecdote he says is found in a Chinese tract on painting:

> The story comes from China, and tells of an old painter who invited friends to see his newest picture. This picture showed a park and a narrow footpath that ran along a stream and through a grove of trees, culminating at the door of a little cottage in the background. When the painter's friends, however, looked around for the painter, they saw that he was gone—that he was in the picture. There, he followed the little path that led to the door, paused before it quite still, turned, smiled, and disappeared through the narrow opening. In the same way, I too, when occupied with my paintpots and brushes, would be suddenly displaced into the picture. I would resemble the porcelain which I had entered in a cloud of colors.*

Here the primary experience of the picture is not an idea (*Vorstellung*) of the picture that originates with a subject, but a mimetic distortion (*Entstellung*) that moves into the picture—a contemplative emptying of the subject. The viewer empties *himself*, subjectlessly entering the picture that can itself open up in this way because it is inspired and inhabited by no one—because it is an image of absence.

*Walter Benjamin, *Berlin Childhood around 1900*, trans. Howard Eiland (Cambridge, MA: Harvard University Press, 2006), 134–135.

It cannot be taught to someone not endowed with it by nature, as can be done with mathematics in which the pupil takes in as much as the master gives out. It cannot be copied as can writing, in which the copy has as much worth as the original. It cannot be reproduced as can sculpture, in which the cast shares with the original the essential merits of the piece. It cannot produce infinite offspring, like printed books. Painting alone retains its nobility, bringing honours singularly to its author and remaining precious and unique. It never gives rise to offspring equal to itself, and such singularity gives it greater excellence than those things that are spread abroad. (Martin Kemp, ed., *Leonardo on Painting*, selected and trans. Martin Kemp and Margaret Walker [New Haven: Yale University Press, 1989], 19)

12. Fu Shen and Jan Stuart, *Challenging the Past: The Paintings of Chang Dai-chien* (Washington: University of Washington Press, 1991), 37.

XIAN ZHAN: SEALS OF LEISURE

1. See Unverzagt, *Der Wandlungsleib*, 186: "The form of the work that accrues over the course of the collection's history reflects the fact that the artistic value of a picture has a social component. ... As the age of a picture grows, so does the power of the social relations made manifest in it."
2. In the examination to become an official, candidates were also required to compose poems on given themes. Huizong, the art-loving Song dynasty emperor, even introduced painting as an examination subject.
3. Signatures on pictures in Europe only came into use from the sixteenth century onward.
4. The picture apparently depicts a betrothal scene.

FUZHI: COPY

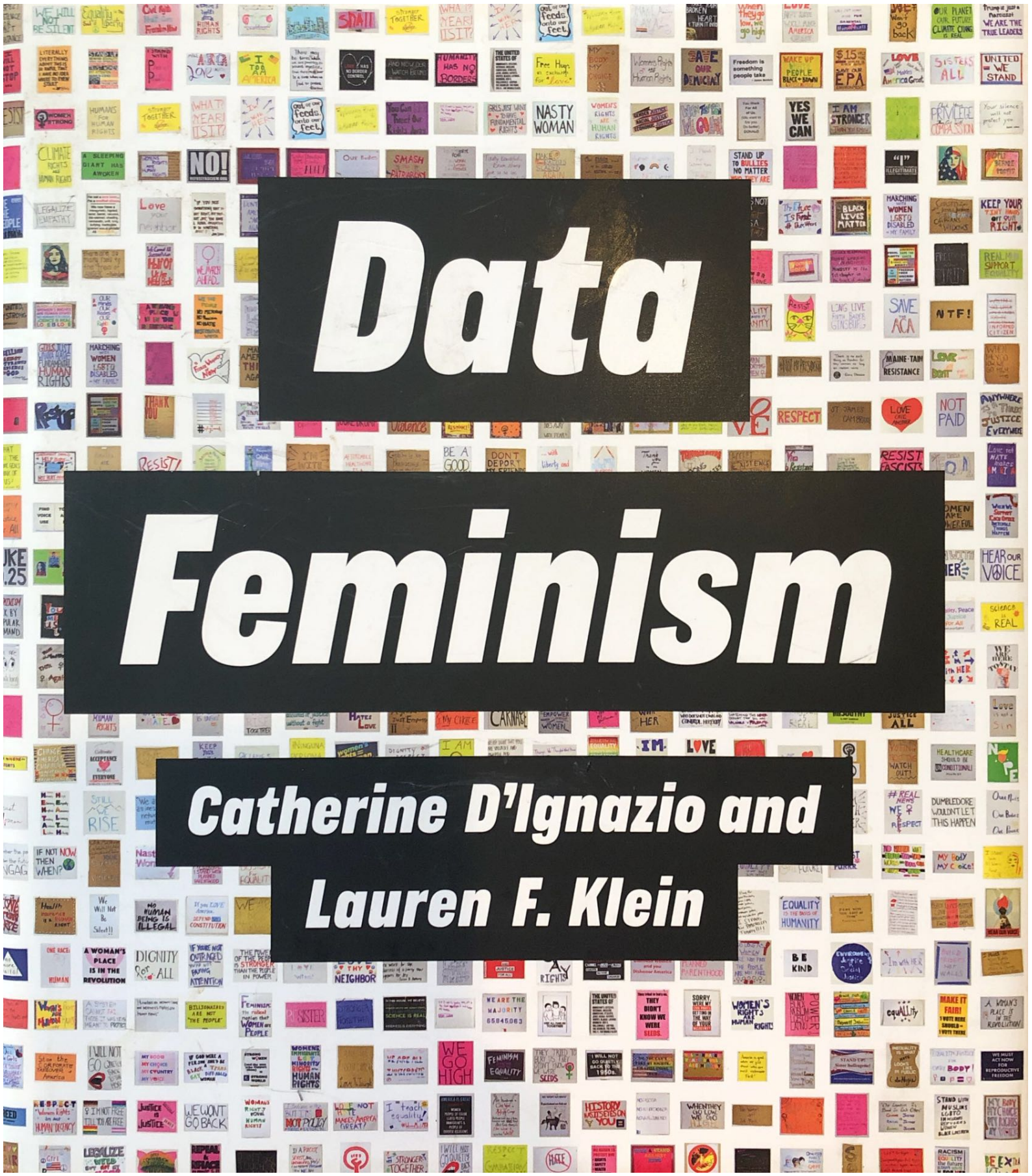
1. Cf. Byung-Chul Han, "Das Klonen und der Ferne Osten," *Lettre International* 64 (2004): 108–109.

2. Lothar Ledderose, *Ten Thousand Things: Module and Mass Production in Chinese Art* (Princeton, NJ: Princeton University Press, 2000), 7.
3. Ibid., 213.

SHANZHAI: FAKE

1. The creativity on which the *shanzhai* movement is based assumes an active process of adaptation and playful combination. This form of creativity cannot be expressed in such hackneyed Asianisms as "not-doing" or contemplation. In his discussion of creativity, even Hans Lenk is unable to transcend these Asianisms:

> In Taoism, for example, if we think of Laozi's *Tao Te Ching*, inaction or *wu wei* plays a highly important role. Creative thought does not take place as a thing compelled or enforced, not when one wants to generate or even force it, but one must tune into letting it happen. *Wu chi* means "no knowledge." Accordingly it means that one does not force the activation of knowledge, but instead adopts a kind of open, primitive, naïve state of comprehension. *Wu yu* is the state of non-desiring which means showing no desires, interests, passion, "disinterested pleasure" in the sense of Kant's *Aesthetics*, or disinterested openness and tolerance. This passive kind of meditation without action, without knowledge, without passions—this is the idea on which creativity is based in Taoist meditation. Letting it happen is considered the mother of creativity. (Hans Lenk, *Kreative Aufstiege: Zur Philosophie und Psychologie der Kreativität* [Frankfurt: Suhrkamp, 2000], 108–109)


Data
Feminism
Catherine D'Ignazio and
Lauren F. Klein

Power and the Matrix of Domination

But first, what do we mean by *power*? We use the term *power* to describe the current configuration of structural privilege and structural oppression, in which some groups experience unearned advantages—because various systems have been designed by people like them and work for people them—and other groups experience systematic disadvantages—because those same systems were not designed by them or with people like them in mind. These mechanisms are complicated, and there are "few pure victims and oppressors," notes influential sociologist Patricia Hill Collins. In her landmark text, *Black Feminist Thought*, first published in 1990, Collins proposes the concept of the *matrix of domination* to explain how systems of power are configured and experienced.[13] It consists of four domains: the structural, the disciplinary, the hegemonic, and the interpersonal. Her emphasis is on the intersection of gender and race, but she makes clear that other dimensions of identity (sexuality, geography, ability, etc.) also result in unjust oppression, or unearned privilege, that become apparent across the same four domains.

The *structural domain* is the arena of laws and policies, along with schools and institutions that implement them. This domain organizes and codifies oppression. Take, for example, the history of voting rights in the United States. The US Constitution did not originally specify who was authorized to vote, so various states had different policies that reflected their local politics. Most had to do with owning property, which,

Table 1.1
The four domains of the matrix of domination

Structural domain Organizes oppression: laws and policies.	**Disciplinary domain** Administers and manages oppression. Implements and enforces laws and policies.
Hegemonic domain Circulates oppressive ideas: culture and media.	**Interpersonal domain** Individual experiences of oppression.

Chart based on concepts introduced by Patricia Hill Collins in *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment.*

conveniently, only men could do. But with the passage of the Fourteenth Amendment in 1868, which granted the rights of US citizenship to those who had been enslaved, the nature of those rights—including voting—were required to be spelled out at the national level for the first time. More specifically, voting was defined as a right reserved for "male citizens." This is a clear instance of codified oppression in the structural domain.

It would take until the passage of the Nineteenth Amendment in 1920 for most (but not all) women to be granted the right to vote.[14] Even still, many state voting laws continued to include literacy tests, residency requirements, and other ways to indirectly exclude people who were not property-owning white men. These restrictions persist today, in the form of practices like dropping names from voter rolls, requiring photo IDs, and limits to early voting—the burdens of which are felt disproportionately by low-income people, people of color, and others who lack the time or resources to jump through these additional bureaucratic hoops.[15] This is the *disciplinary domain* that Collins names: the domain that administers and manages oppression through bureaucracy and hierarchy, rather than through laws that explicitly encode inequality on the basis of someone's identity.[16]

Neither of these domains would be possible without *the hegemonic domain*, which deals with the realm of culture, media, and ideas. Discriminatory policies and practices in voting can only be enacted in a world that already circulates oppressive ideas about, for example, who counts as a citizen in the first place. Consider an anti-suffragist pamphlet from the 1910s that proclaims, "You do not need a ballot to clean out your sink spout."[17] Pamphlets like these, designed to be literally passed from hand to hand, reinforced preexisting societal views about the place of women in society. Today, we have animated GIFs instead of paper pamphlets, but the hegemonic function is the same: to consolidate ideas about who is entitled to exercise power and who is not.

The final part of the matrix of domination is the *interpersonal domain*, which influences the everyday experience of individuals in the world. How would you feel if you were a woman who read that pamphlet, for example? Would it have more or less of an impact if a male family member gave it to you? Or, for a more recent example, how would you feel if you took time off from your hourly job to go cast your vote, only to discover when you got there that your name had been purged from the official voting roll or that there was a line so long that it would require that you miss half a day's pay, or stand for hours in the cold, or ... the list could go on. These are examples of how it *feels* to know that systems of power are not on your side and, at times, are actively seeking to take away the small amount of power that you do possess.[18]

The matrix of domination works to uphold the undue privilege of *dominant* groups while unfairly oppressing *minoritized* groups. What does this mean? Beginning in this chapter and continuing throughout the book, we use the term *minoritized* to describe groups of people who are positioned in opposition to a more powerful social group. While the term *minority* describes a social group that is comprised of fewer people, *minoritized* indicates that a social group is actively devalued and oppressed by a dominant group, one that holds more economic, social, and political power. With respect to gender, for example, men constitute the dominant group, while all other genders constitute minoritized groups. This remains true even as women actually constitute a majority of the world population. *Sexism* is the term that names this form of oppression. In relation to race, white people constitute the dominant group (racism); in relation to class, wealthy and educated people constitute the dominant group (classism); and so on.[19]

Using the concept of the matrix of domination and the distinction between dominant and minoritized groups, we can begin to examine how power unfolds in and around data. This often means asking uncomfortable questions: who is doing the work of data science (and who is not)? Whose goals are prioritized in data science (and whose are not)? And who benefits from data science (and who is either overlooked or actively harmed)?[20] These questions are uncomfortable because they unmask the inconvenient truth that there are groups of people who are disproportionately benefitting from data science, and there are groups of people who are disproportionately harmed. Asking these *who questions* allows us, as data scientists ourselves, to start to see how privilege is baked into our data practices and our data products.[21]

Data Science by Whom?

It is important to acknowledge the elephant in the server room: the demographics of data science (and related occupations like software engineering and artificial

intelligence research) do not represent the population as a whole. According to the most recent data from the US Bureau of Labor Statistics, released in 2018, only 26 percent of those in "computer and mathematical occupations" are women.[22] And across all of those women, only 12 percent are Black or Latinx women, even though Black and Latinx women make up 22.5 percent of the US population.[23] A report by the research group AI Now about the diversity crisis in artificial intelligence notes that women comprise only 15 percent of AI research staff at Facebook and 10 percent at Google.[24] These numbers are probably not a surprise. The more surprising thing is that those numbers are getting worse, not better. According to a research report published by the American Association of University Women in 2015, women computer science graduates in the United States peaked in the mid-1980s at 37 percent, and we have seen a steady decline in the years since then to 26 percent today (figure 1.2).[25] As "data analysts" (low-status number crunchers) have become rebranded as "data scientists" (high status

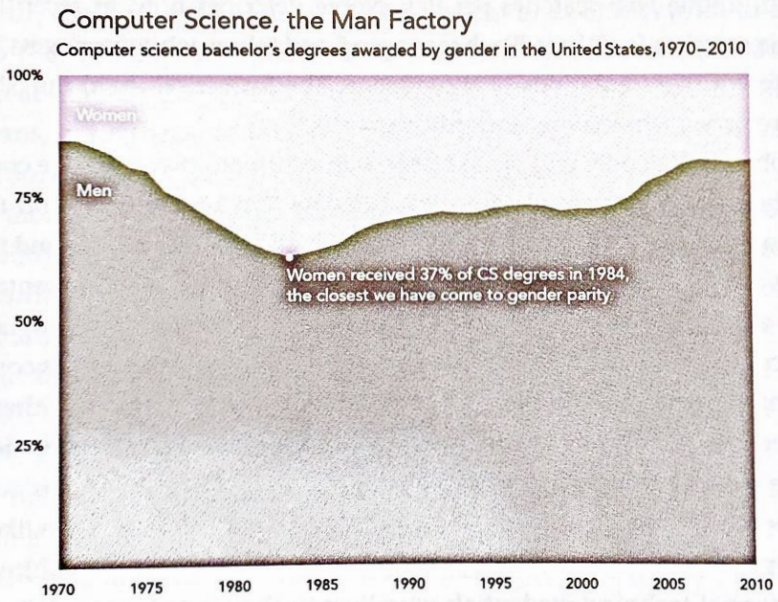


Figure 1.2
Computer science has always been dominated by men and the situation is worsening (even while many other scientific and technical fields have made significant strides toward gender parity). Women awarded bachelor's degrees in computer science in the United States peaked in the mid-1980s at 37 percent, and we have seen a steady increase in the ratio of men to women in the years since then. This particular report treated gender as a binary, so there was no data about nonbinary people. Graphic by Catherine D'Ignazio. Data from the National Center for Education Statistics.

researchers), women are being pushed out in order to make room for more highly valued and more highly compensated men.[26]

There are not disparities only along gender lines in the higher education pipeline. The same report noted specific underrepresentation for Native American women, multiracial women, white women, and all Black and Latinx people. So is it really a surprise that each day brings a new example of data science being used to disempower and oppress minoritized groups? In 2018, it was revealed that Amazon had been developing an algorithm to screen its first-round job applicants. But because the model had been trained on the resumes of prior applicants, who were predominantly male, it developed an even stronger preference for male applicants. It downgraded resumes with the word *women* and graduates of women's colleges. Ultimately, Amazon had to cancel the project.[27] This example reinforces the work of Safiya Umoja Noble, whose book, *Algorithms of Oppression*, has shown how both gender and racial biases are encoded into some of the most pervasive data-driven systems—including Google search, which boasts over five billion unique web searches per day. Noble describes how, as recently as 2016, comparable searches for "three Black teenagers" and "three white teenagers" turned up wildly different representations of those teens. The former returned mugshots, while the latter returned wholesome stock photography.[28]

The problems of gender and racial bias in our information systems are complex, but some of their key causes are plain as day: the data that shape them, and the models designed to put those data to use, are created by small groups of people and then scaled up to users around the globe. But those small groups are not at all representative of the globe as a whole, nor even of a single city in the United States. When data teams are primarily composed of people from dominant groups, those perspectives come to exert outsized influence on the decisions being made—to the exclusion of other identities and perspectives. This is not usually intentional; it comes from the ignorance of being on top. We describe this deficiency as a *privilege hazard*.

How does this come to pass? Let's take a minute to imagine what life is like for someone who epitomizes the dominant group in data science: a straight, white, cisgender man with formal technical credentials who lives in the United States. When he looks for a home or applies for a credit card, people are eager for his business. People smile when he holds his girlfriend's hand in public. His body doesn't change due to childbirth or breastfeeding, so he does not need to think about workplace accommodations. He presents his social security number in jobs as a formality, but it never hinders his application from being processed or brings him unwanted attention. The ease with which he traverses the world is invisible to him because it has been designed for people

just like him. He does not think about how life might be different for everyone else. In fact, it is difficult for him to imagine that at all.

This is the *privilege hazard*: the phenomenon that makes those who occupy the most privileged positions among us—those with good educations, respected credentials, and professional accolades—so poorly equipped to recognize instances of oppression in the world.[29] They lack what Anita Gurumurthy, executive director of IT for Change, has called "the empiricism of lived experience."[30] And this lack of lived experience—this evidence of how things truly *are*—profoundly limits their ability to foresee and prevent harm, to identify existing problems in the world, and to imagine possible solutions.

The privilege hazard occurs at the level of the individual—in the interpersonal domain of the matrix of domination—but it is much more harmful in aggregate because it reaches the hegemonic, disciplinary and structural domains as well. So it matters deeply that data science and artificial intelligence are dominated by elite white men because it means there is a collective privilege hazard so great that it would be a profound surprise if they could actually identify instances of bias prior to unleashing them onto the world. Social scientist Kate Crawford has advanced the idea that the biggest threat from artificial intelligence systems is not that they will become smarter than humans, but rather that they will hard-code sexism, racism, and other forms of discrimination into the digital infrastructure of our societies.[31]

What's more, the same cis het white men responsible for designing those systems lack the ability to detect harms and biases in their systems once they've been released into the world.[32] In the case of the "three teenagers" Google searches, for example, it was a young Black teenager that pointed out the problem and a Black scholar who wrote about the problem. The burden consistently falls upon those more intimately familiar with the privilege hazard—in data science as in life—to call out the creators of those systems for their limitations.

For example, Joy Buolamwini, a Ghanaian-American graduate student at MIT, was working on a class project using facial-analysis software.[33] But there was a problem—the software couldn't "see" Buolamwini's dark-skinned face (where "seeing" means that it detected a face in the image, like when a phone camera draws a square around a person's face in the frame). It had no problem seeing her lighter-skinned collaborators. She tried drawing a face on her hand and putting it in front of the camera; it detected that. Finally, Buolamwini put on a white mask, essentially going in "whiteface" (figure 1.3).[34] The system detected the mask's facial features perfectly.

Digging deeper into the code and benchmarking data behind these systems, Buolamwini discovered that the dataset on which many of facial-recognition algorithms are tested contains 78 percent male faces and 84 percent white faces. When she did

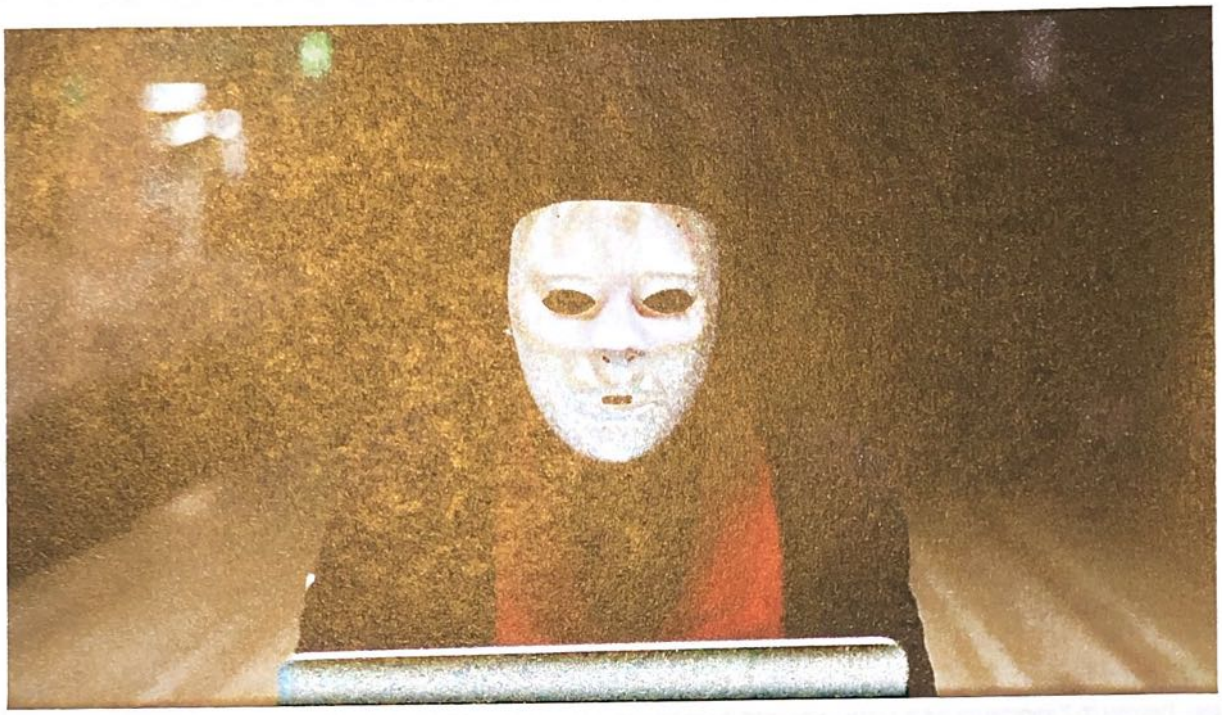


Figure 1.3
Joy Buolamwini found that she had to put on a white mask for the facial detection program to "see" her face. Buolamwini is now founder of the Algorithmic Justice League. Courtesy of Joy Buolamwini.

an intersectional breakdown of another test dataset—looking at gender and skin type together—only 4 percent of the faces in that dataset were women and dark-skinned. In their evaluation of three commercial systems, Buolamwini and computer scientist Timnit Gebru showed that darker-skinned women were up to forty-four times more likely to be misclassified than lighter-skinned males.[35] It's no wonder that the software failed to detect Buolamwini's face: both the training data and the benchmarking data relegate women of color to a tiny fraction of the overall dataset.[36]

This is the privilege hazard in action—that no coder, tester, or user of the software had previously identified such a problem or even thought to look. Buolamwini's work has been widely covered by the national media (by the *New York Times*, by CNN, by the *Economist*, by *Bloomberg BusinessWeek*, and others) in articles that typically contain a hint of shock.[37] This is a testament to the social, political, and technical importance of the work, as well as to how those in positions of power—not just in the field of data science, but in the mainstream media, in elected government, and at the heads of corporations—are so often surprised to learn that their "intelligent technologies" are not so intelligent after all. (They need to read data journalist Meredith Broussard's book *Artificial Unintelligence*).[38] For another example, think back to the introduction of this

book, where we quoted Shetterly as reporting that Christine Darden's white male manager was "shocked at the disparity" between the promotion rates of men and women. We can speculate that Darden herself wasn't shocked, just as Buolamwini and Gebru likely were not entirely shocked at the outcome of their study either. When sexism, racism, and other forms of oppression are publicly unmasked, it is almost never surprising to those who experience them.

For people in positions of power and privilege, issues of race and gender and class and ability—to name only a few—are OPP: other people's problems. Author and antiracist educator Robin DiAngelo describes instances like the "shock" of Darden's boss or the surprise in the media coverage of Buolamwini's various projects as a symptom of the "racial innocence" of white people.[39] In other words, those who occupy positions of privilege in society are able to remain innocent of that privilege. Race becomes something that only people of color have. Gender becomes something that only women and nonbinary people have. Sexual orientation becomes something that all people *except* heterosexual people have. And so on. A personal anecdote might help illustrate this point. When we published the first draft of this book online, Catherine told a colleague about it. His earnestly enthusiastic response was, "Oh great! I'll show it to my female graduate students!" To which Catherine rejoined, "You might want to show it to your other students, too."

If things were different—if the 79 percent of engineers at Google who are male were specifically trained in structural oppression before building their data systems (as social workers are before they undertake social work)—then their overrepresentation might be very slightly less of a problem.[40] But in the meantime, the onus falls on the individuals who already feel the adverse effects of those systems of power to prove, over and over again, that racism and sexism exist—in datasets, in data systems, and in data science, as in everywhere else.

Buolamwini and Gebru identified how pale and male faces were overrepresented in facial detection training data. Could we just fix this problem by diversifying the data set? One solution to the problem would appear to be straightforward: create a more representative set of training and benchmarking data for facial detection models. In fact, tech companies are starting to do exactly this. In January 2019, IBM released a database of one million faces called Diversity in Faces (DiF).[41] In another example, journalist Amy Hawkins details how CloudWalk, a startup in China in need of more images of faces of people of African descent, signed a deal with the Zimbabwean government for it to provide the images the company was lacking.[42] In return for sharing its data, Zimbabwe will receive a national facial database and "smart" surveillance infrastructure that it can install in airports, railways, and bus stations.

It might sound like an even exchange, but Zimbabwe has a dismal record on human rights. Making things worse, CloudWalk provides facial recognition technologies to the Chinese police—a conflict of interest so great that the global nonprofit Human Rights Watch voiced its concern about the deal.[43] Face harvesting is happening in the US as well. Researchers Os Keyes, Nikki Stevens and Jacqueline Wernimont have shown how immigrants, abused children, and dead people are some of the groups whose faces have been used to train software—without their consent.[44] So is a diverse database of faces really a good idea? Voicing his concerns in response to the announcement of Buolamwini and Gebru's 2018 study on Twitter, an Indigenous Marine veteran shot back, "I hope facial recognition software has a problem identifying my face too. That'd come in handy when the police come rolling around with their facial recognition truck at peaceful demonstrations of dissent, cataloging all dissenters for 'safety and security.'"[45]

Better detection of faces of color cannot be characterized as an unqualified good. More often than not, it is enlisted in the service of increased oppression, greater surveillance, and targeted violence. Buolamwini understands these potential harms and has developed an approach that works across all four domains of the matrix of domination to address the underlying issues of power that are playing out in facial analysis technology. Buolamwini and Gebru first quantified the disparities in the dataset—a technical audit, which falls in the disciplinary domain of the matrix of domination. Then, Buolamwini went on to launch the Algorithmic Justice League, an organization that works to highlight and intervene in instances of algorithmic bias. On behalf of the AJL, Buolamwini has produced viral poetry projects and given TED talks—taking action in the hegemonic domain, the realm of culture and ideas. She has advised on legislation and professional standards for the field of computer vision and called for a moratorium on facial analysis in policing on national media and in Congress.[46] These are actions operating in the structural domain of the matrix of domination—the realm of law and policy. Throughout these efforts, the AJL works with students and researchers to help guide and shape their own work—the interpersonal domain. Taken together, Buolamwini's various initiatives demonstrate how any "solution" to bias in algorithms and datasets must tackle more than technical limitations. In addition, they present a compelling model for the data scientist as public intellectual—who, yes, works on technical audits and fixes, but also works on cultural, legal, and political efforts too.

While equitable representation—in datasets and data science workforces—is important, it remains window dressing if we don't also transform the institutions that produce and reproduce those biased outcomes in the first place. As doctoral health student

Arrianna Planey, quoting Robert M. Young, states, "A racist society will give you a racist science."[47] We cannot filter out the downstream effects of sexism and racism without also addressing their root cause.

Figure 1.4
The Library of Missing Datasets, by Mimi Onuoha (2016) is a list of datasets that are not collected because of bias, lack of social and political will, and structural disregard. Courtesy of Mimi Onuoha. Photo by Brandon Schulman.

As goods and services become commoditized, the customer experiences that companies create will matter most.

WELCOME TO THE EXPERIENCE ECONOMY

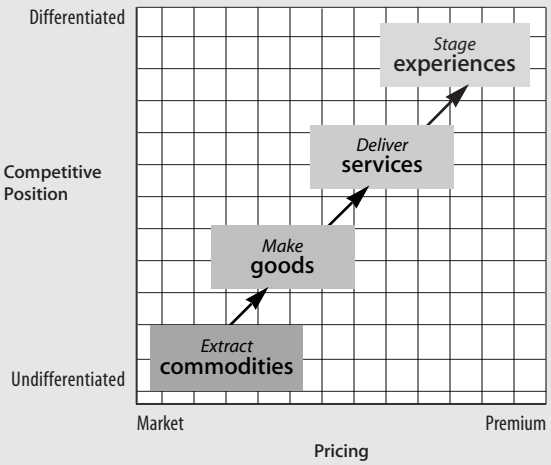
BY B. JOSEPH PINE II AND JAMES H. GILMORE

How do economies change? The entire history of economic progress can be recapitulated in the four-stage evolution of the birthday cake. As a vestige of the agrarian economy, mothers made birthday cakes from scratch, mixing farm commodities (flour, sugar, butter, and eggs) that together cost mere dimes. As the goods-based industrial economy advanced, moms paid a dollar or two to Betty Crocker for premixed ingredients. Later, when the service economy took hold, busy parents ordered cakes from the bakery or grocery store, which, at $10 or $15, cost ten times as much as the packaged ingredients. Now, in the time-starved 1990s, parents neither make the birthday cake nor even throw the party. Instead, they spend $100 or more to "outsource" the entire event to Chuck E. Cheese's, the Discovery Zone, the Mining Company, or some other business that stages a memorable event for the kids – and often throws in the cake for free. Welcome to the emerging *experience economy*.

Economists have typically lumped experiences in with services, but experiences are a distinct economic offering, as different from services as services are from goods. Today we can identify and describe this fourth economic offering because consumers unquestionably desire experiences, and more and more businesses are responding by explicitly designing and promoting them. As services, like goods before them, increasingly become commoditized – think of long-distance telephone services sold solely on price – experiences have emerged as the next step in what we call the *progression of economic value*. (See the exhibit "The Progression of Economic Value.") From now on, leading-edge companies – whether they sell to consumers or businesses – will find

B. Joseph Pine II *and* ***James H. Gilmore*** *are cofounders of Strategic Horizons LLP, based in Cleveland, Ohio. They are coauthors of* The Experience Economy: Work Is Theatre and Every Business a Stage, *to be published by the Harvard Business School Press in April 1999. They are the authors of "The Four Faces of Mass Customization" (HBR January–February 1997) and can be reached at pine@gilmore@customization.com.*

The Progression of Economic Value



that the next competitive battleground lies in staging experiences.

An experience is not an amorphous construct; it is as real an offering as any service, good, or commodity. In today's service economy, many companies simply wrap experiences around their traditional offerings to sell them better. To realize the full benefit of staging experiences, however, businesses must deliberately design engaging experiences that command a fee. This transition from selling services to selling experiences will be no easier for established companies to undertake and weather than the last great economic shift, from the industrial to the service economy. Unless companies want to be in a commoditized business, however, they will be compelled to upgrade their offerings to the next stage of economic value.

The question, then, isn't whether, but when – and how – to enter the emerging experience economy. An early look at the characteristics of experiences and the design principles of pioneering experience stagers suggests how companies can begin to answer this question.

Staging Experiences that Sell

To appreciate the difference between services and experiences, recall the episode of the old television show *Taxi* in which Iggy, a usually atrocious (but fun-loving) cab driver, decided to become the best taxi driver in the world. He served sandwiches and drinks, conducted tours of the city, and even sang Frank Sinatra tunes. By engaging passengers in a way that turned an ordinary cab ride into a memorable event, Iggy created something else entirely – a distinct economic offering. The *experience* of riding in his cab was more valuable to his customers than the service of being transported by the cab – and in the TV show, at least, Iggy's customers happily responded by giving bigger tips. By asking to go around the block again, one patron even *paid more* for poorer service just to prolong his enjoyment. The service Iggy provided – taxi transportation – was simply the stage for the experience that he was really selling.

An experience occurs when a company intentionally uses services as the stage, and goods as props, to engage individual customers in a way that creates a memorable event. Commodities are fungible, goods tangible, services intangible, and experiences *memorable.* (See the chart "Economic Distinc-

Economic Distinctions

Economic Offering	Commodities	Goods	Services	Experiences
Economy	Agrarian	Industrial	Service	Experience
Economic Function	Extract	Make	Deliver	Stage
Nature of Offering	Fungible	Tangible	Intangible	Memorable
Key Attribute	Natural	Standardized	Customized	Personal
Method of Supply	Stored in bulk	Inventoried after production	Delivered on demand	Revealed over a duration
Seller	Trader	Manufacturer	Provider	Stager
Buyer	Market	User	Client	Guest
Factors of Demand	Characteristics	Features	Benefits	Sensations

tions.") Buyers of experiences – we'll follow the lead of experience-economy pioneer Walt Disney and call them "guests" – value what the company reveals over a duration of time. While prior economic offerings – commodities, goods, and services – are external to the buyer, experiences are inherently personal, existing only in the mind of an individual who has been engaged on an emotional, physical, intellectual, or even spiritual level. Thus, no two people can have the same experience, because each experience derives from the interaction between the staged event (like a theatrical play) and the individual's state of mind.

Experiences have always been at the heart of the entertainment business – a fact that Walt Disney and the company he founded have creatively exploited. But today the concept of selling an entertainment experience is taking root in businesses far removed from theaters and amusement parks. New technologies, in particular, encourage whole new genres of experience, such as interactive games, Internet chat rooms and multiplayer games, motion-based simulators, and virtual reality. The growing processing power required to render ever-more immersive experiences now drives demand for the goods and services of the computer industry. In a speech made at the November 1996 COMDEX computer trade show, Intel chairman Andrew Grove declared, "We need to look at our business as more than simply the building and selling of personal computers. Our business is the delivery of information and lifelike interactive experiences."

At theme restaurants such as the Hard Rock Cafe, Planet Hollywood, or the House of Blues, the food is just a prop for what's known as "eatertainment." And stores such as Niketown, Cabella's, and Recreational Equipment Incorporated draw consumers in by offering fun activities, fascinating displays, and promotional events (sometimes labeled "shoppertainment" or "entertailing").

But experiences are not exclusively about entertainment; companies stage an experience whenever they engage customers in a personal, memorable way. In the travel business, former British Airways chairman Sir Colin Marshall has noted that the "commodity mind-set" is to "think that a business is merely performing a function – in our case, transporting people from point A to point B on time and at the lowest possible price." What British Airways does, according to Sir Colin, is "to go beyond the function and compete on the basis of providing an experience." (See "Competing on Customer Service: An Interview with British Airways' Sir Colin Marshall," HBR November–December 1995.) The company uses its base service (the travel itself) as the stage for a distinctive en route experience – one that attempts to transform air travel into a respite from the traveler's normally frenetic life.

Neither are experiences only for consumer industries. Companies consist of people, and business-to-business settings also present stages for experiences. For example, a Minneapolis computer-installation and repair company calls itself the Geek Squad. Its "special agents" costume themselves in white shirts with thin black ties and pocket protectors, carry badges, drive old cars, and turn a humdrum activity into a memorable encounter. Similarly, many companies hire theater troupes – like the St. Louis-based trainers One World Music, facilitators of a program called Synergy through Samba – to turn otherwise ordinary meetings into improvisational events that encourage breakthrough thinking.

Today the concept of selling experiences is spreading beyond theaters and theme parks.

Business-to-business marketers increasingly create venues as elaborate as any Disney attraction in which to sell their goods and services. In June 1996, Silicon Graphics, for example, opened its Visionarium Reality Center at corporate headquarters in Mountain View, California, to bring customers and engineers together in an environment where they can interact with real-time, three-dimensional product visualizations. Customers can view, hear, and touch – as well as drive, walk, or fly – through myriad product possibilities. "This is experiential computing at its ultimate, where our customers can know what their products will look like, sound like, feel like before manufacturing," said then chairman and CEO Edward McCracken.

You Are What You Charge For

Notice, however, that while all of these companies stage experiences, most are still charging for their goods and services. Companies generally move from one economic stage to the next in incremental steps. In its heyday in the 1960s and 1970s, IBM's slogan was "IBM Means Service," and the computer manufacturer indeed lavished services – for free –

on any company that would buy its hardware goods. It planned facilities, programmed code, integrated other companies' equipment, and repaired its own machines; its service offerings overwhelmed the competition. But eventually IBM had to charge customers for what it had been giving away for free, when a Justice Department suit required the company to unbundle its hardware and software. The government order notwithstanding, IBM couldn't afford to continue to meet increasing customer-service demands without explicitly charging for them. Services, it turned out, were the company's most valued offerings. Today, with its mainframe computers long since commoditized, IBM's Global Services unit grows at double-digit annual rates. The company no longer gives away its services to sell its goods. In fact, the deal is reversed: the company will buy its clients' hardware if they'll contract with Global Services to manage their information systems. IBM still manufactures computers, but now it's in the business of providing services.

Companies should think about what they would do differently if they charged admission.

It's an indication of the maturity of the service economy that IBM and other manufacturers now make greater profits from the services than from the goods they provide. General Electric's GE Capital unit and the financial arms of the Big Three automakers are cases in point. Likewise, it's an indication of the immaturity of the experience economy that most companies providing experiences – like the Hard Rock Cafe, the Geek Squad, or Silicon Graphics – don't yet explicitly charge for the events that they stage.

No company sells experiences as its economic offering unless it actually charges guests an admission fee. An event created just to increase customer preference for the commoditized goods or services that a company actually sells is not an economic offering. But even if a company rejects (for now) charging admission to events that it stages, its managers should already be asking themselves what they would do differently if they were to charge admission. The answers will help them see how their company might begin to move forward into the experience economy, for such an approach demands the design of richer experiences.

Movie theaters already charge admission to see featured films, but Jim Loeks, part-owner of the Star theater complex in Southfield, Michigan, told *Forbes* magazine that "it should be worth the price of the movie just to go into the theater." Star charges 3 million customers a year 25% higher admission for a movie than a local competitor does because of the fun-house experience it provides. Soon, perhaps, with 65,000 square feet of restaurants and stores being added to the complex, Star will charge its customers admission just to get into the complex.

Some retailers already border on the experiential. At the Sharper Image or Brookstone, notice how many people play with the gadgets, listen to miniaturized stereo equipment, sit in massage chairs, and then leave without paying for what they valued, namely, the experience. Could these stores charge admission? Not as they are currently managed. But if they did charge an admission fee, they would be forced to stage a much better experience to attract paying guests. The merchandise mix would need to change more often – daily or even hourly. The stores would have to add demonstrations, showcases, contests, and other attractions to enhance the customer experience.

With its Niketown stores, Nike is almost in the experience business. To avoid alienating its existing retail channels, Nike created Niketown as a merchandising exposition. It's ostensibly for show – to build the brand image and stimulate buying at other retail outlets – not for selling. If that is so, then why not explicitly charge customers for experiencing Niketown? Would people pay? People have already queued to enter the Niketown on Chicago's Michigan Avenue. An admission fee would force Nike to stage more engaging events inside. The stores might actually use the basketball court, say, to stage one-on-one games or rounds of horse with National Basketball Association players. Afterward customers could buy customized Nike T-shirts, commemorating the date and score of events – complete with an action photo of the winning hoop. There might be more interactive kiosks for educational exploration of past athletic events. Virtual reality machines could let you, as Nike's advertising attests, be Tiger Woods. Nike could probably generate as much admission-based revenue per square foot from Niketown as the Walt Disney Company does from its entertainment venues – and as Disney *should* (but does not) yield from its own retail stores. For the premier company of the experience economy, Disney's specialty re-

tailing outside of its own theme parks disappoints. Its mall stores aren't much different from anyone else's, precisely because Disney doesn't charge admission to them – and so doesn't bother creating the extraordinary experiences it so expertly creates elsewhere.

An entrepreneur in Israel has entered the experience economy with the opening of Cafe Ke'ilu, which roughly translates as "Cafe Make Believe." Manager Nir Caspi told a reporter that people come to cafés to be seen and to meet people, not for the food; Cafe Ke'ilu pursues that observation to its logical conclusion. The establishment serves its customers plates and mugs that are empty and charges guests $3 during the week and $6 on weekends for the social experience.

Charging admission – requiring customers to pay for the experience – does not mean that companies have to stop selling goods and services. Disney generates significant profits from parking, food, and other service fees at its theme parks as well as from the sale of memorabilia. But without the staged experiences of the company's theme parks, cartoons, movies, and TV shows, customers would have nothing to remember – and Disney would have no characters to exploit.

In the full-fledged experience economy, retail stores and even entire shopping malls will charge admission before they let a consumer even set foot in them. Some shopping malls, in fact, already do charge admission. We're not thinking of the Mall of America outside of Minneapolis, which contains an amusement park; it charges for the rides, but the shopping is still free. We're referring to the Gilroy Garlic Festival in California, the Minnesota Renaissance Festival, the Kitchener-Waterloo Oktoberfest in Ontario, Canada, and other seasonal festivals that are really outdoor shopping malls and do indeed charge admission. Consumers judge them worth the fees because the festival operators script distinctive experiences around enticing themes, as well as stage activities that captivate customers before, after, and while they shop. With nearly every customer leaving with at least one bag of merchandise, these festival experiences clearly capture shopping dollars that otherwise would be spent at traditional malls and retail outlets.

The business equivalent of a shopping mall is a trade show – a place for finding, learning about, and, if a need is met, purchasing exhibitors' offerings. Trade-show operators already charge admission to the experiences they create; individual business-to-business companies will need to do the same, essentially charging customers to sell to them. Diamond Technology Partners for instance, stages the Diamond Exchange, a series of forums that help members explore the digital future. Current and potential clients pay tens of thousands of dollars annually to attend because what they gain – fresh insights, self-discovery, and engaging interactions – is worth it. No one minds that in staging the event, Diamond greatly improves its chances of selling follow-up consulting work.

The Characteristics of Experiences

Before a company can charge admission, it must design an experience that customers judge to be worth the price. Excellent design, marketing, and delivery will be every bit as crucial for experiences as they are for goods and services. Ingenuity and innovation will always precede growth in revenue. Yet experiences, like goods and services, have their own distinct qualities and characteristics and present their own design challenges.

One way to think about experiences is across two dimensions. The first corresponds to *customer participation.* At one end of the spectrum lies passive participation, in which customers don't affect the performance at all. Such participants include symphony-goers, for example, who experience the event as observers or listeners. At the other end of the spectrum lies active participation, in which customers play key roles in creating the performance or event that yields the experience. These participants include skiers. But even people who turn out to watch a ski race are not completely passive participants; simply by being there, they contribute to the visual and aural event that others experience.

> Some companies will eventually be like trade shows, charging customers to sell to them.

The second dimension of experience describes the *connection,* or environmental relationship, that unites customers with the event or performance. At one end of the connection spectrum lies absorption, at the other end, immersion. People viewing the Kentucky Derby from the grandstand can absorb the event taking place beneath and in front of them; meanwhile, people standing in the

infield are immersed in the sights, sounds, and smells that surround them. Furiously scribbling notes while listening to a physics lecture is more absorbing than reading a textbook; seeing a film at the theater with an audience, large screen, and stereophonic sound is more immersing than watching the same film on video at home.

We can sort experiences into four broad categories according to where they fall along the spectra of the two dimensions. (See the exhibit "The Four Realms of an Experience.") The kinds of experiences most people think of as entertainment – watching television, attending a concert – tend to be those in which customers participate more passively than actively; their connection with the event is more likely one of absorption than of immersion. Educational events – attending a class, taking a ski lesson – tend to involve more active participation, but students (customers, if you will) are still more outside the event than immersed in the action. Escapist experiences can teach just as well as educational events can, or amuse just as well as entertainment, but they involve greater customer immersion. Acting in a play, playing in an orchestra, or descending the Grand Canyon involve both active participation and immersion in the experience. If you minimize the customers' active participation, however, an escapist event becomes an experience of the fourth kind – the esthetic. Here customers or participants are immersed in an activity or environment, but they themselves have little or no effect on it – like a tourist who merely views the Grand Canyon from its rim or like a visitor to an art gallery.

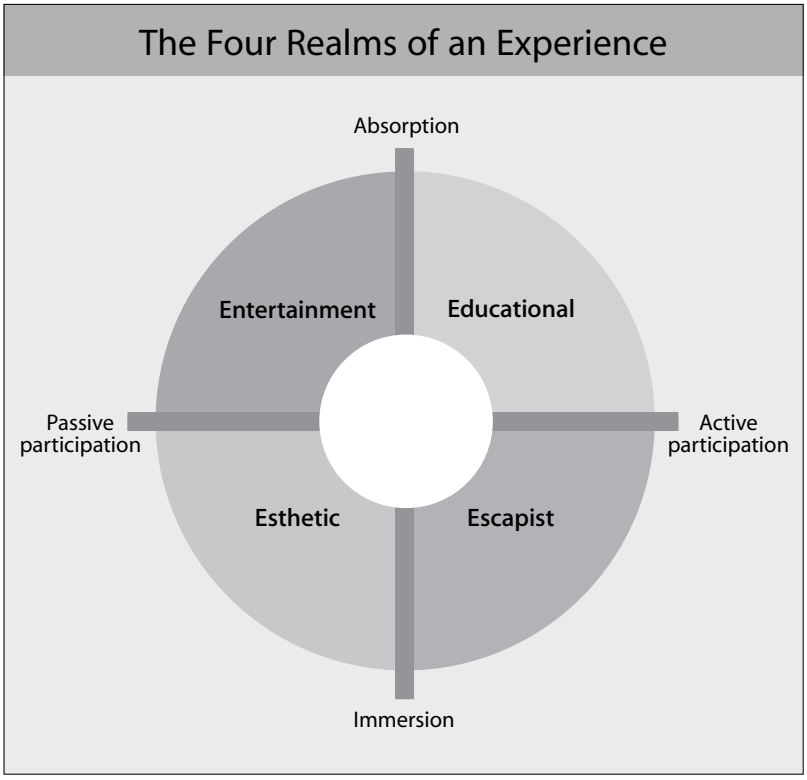
Generally, we find that the richest experiences – such as going to Disney World or gambling in a Las Vegas casino – encompass aspects of all four realms, forming a "sweet spot" around the area where the spectra meet. But still, the universe of possible experiences is vast. Eventually, the most significant question managers can ask themselves is "What specific experience will my company offer?" That experience will come to define their business.

Experiences, like goods and services, have to meet a customer need; they have to work; and they have to be deliverable. Just as goods and services result from an iterative process of research, design, and development, experiences derive from an iterative process of exploration, scripting, and staging – capabilities that aspiring experience merchants will need to master.

Designing Memorable Experiences

We expect that experience design will become as much a business art as product design and process design are today. Indeed, design principles are already apparent from the practices of and results obtained by companies that have (or nearly have) advanced into the experience economy. We have identified five key experience-design principles.

Theme the experience. Just hear the name of any "eatertainment" restaurant – Hard Rock Cafe, Planet Hollywood, or the Rainforest Cafe, to name a few – and you instantly know what to expect when you enter the establishment. The proprietors have taken the first, crucial step in staging an experience by envisioning a well-defined theme. One poorly conceived, on the other hand, gives customers nothing around which to organize the impressions they encounter, and the experience yields no lasting memory. An incoherent theme is like Gertrude Stein's Oakland: "There is no there there." Retailers often offend the principle. They talk of "the shopping experience" but fail to create a theme that ties the disparate merchandising presentations together into a staged experience. Home-appliance and electronics retailers in partic-



ular show little thematic imagination. Row upon row of washers and dryers and wall after wall of refrigerators accentuate the sameness of different companies' stores. Shouldn't there be something distinctive about an establishment called Circuit City, for example?

Consider the Forum Shops in Las Vegas, a mall that displays its distinctive theme – an ancient Roman marketplace – in every detail. The Simon DeBartolo Group, which developed the mall, fulfills this motif through a panoply of architectural effects. These include marble floors, stark white pillars, "outdoor" cafés, living trees, flowing fountains – and even a painted blue sky with fluffy white clouds that yield regularly to simulated storms, complete with lightning and thunder. Every mall entrance and every storefront is an elaborate Roman re-creation. Every hour inside the main entrance, statues of Caesar and other Roman luminaries come to life and speak. "Hail, Caesar!" is a frequent cry, and Roman centurions periodically march through on their way to the adjacent Caesar's Palace casino. The Roman theme even extends into some of the shops. A jewelry store's interior, for instance, features scrolls, tablets, Roman numerals, and gold draperies. The theme implies opulence, and the mall's 1997 sales – more than $1,000 per square foot, compared with a typical mall's sales of less than $300 – suggest that the experience works.

An effective theme is concise and compelling. It is not a corporate mission statement or a marketing tag line. It needn't be publicly articulated in writing. But the theme must drive all the design elements and staged events of the experience toward a unified story line that wholly captivates the customer. Educational Discoveries and Professional Training International of Orem, Utah, stage a day-long course on basic accounting skills to nonfinancial managers. Their exquisitely simple theme – running a lemonade stand – turns learning into an experience. Students use real lemons and lemonade, music, balloons, and a good deal of ballyhoo while they create a corporate financial statement. The theme unifies the experience in the students' minds and helps make the learning memorable.

Harmonize impressions with positive cues. While the theme forms the foundation, the experience must be rendered with indelible impressions. Impressions are the "takeaways" of the experience; they fulfill the theme. To create the desired impressions, companies must introduce cues that affirm the nature of the experience to the guest. Each cue must support the theme, and none should be inconsistent with it.

George Harrop, founder of Barista Brava, a franchised chain of coffee bars based in Washington, D.C., developed the company's theme of "the marriage of Old-World Italian espresso bars with fast-paced American living." The interior decor supports the Old World theme, and the carefully designed pattern of the floor tiles and counters encourages customers to line up without the usual signage or ropes that would detract from that theme. The impressions convey quick service in a soothing setting. Furthermore, Harrop encourages *baristas* to remember faces so that regular customers are handed their usual order without even having to ask.

To create the desired impression, companies must provide cues that affirm the nature of the experience.

Even the smallest cue can aid the creation of a unique experience. When a restaurant host says, "Your table is ready," no particular cue is given. But when a Rainforest Cafe host declares, "Your adventure is about to begin," it sets the stage for something special.

It's the cues that make the impressions that create the experience in the customer's mind. An experience can be unpleasant merely because some architectural feature has been overlooked, underappreciated, or uncoordinated. Unplanned or inconsistent visual and aural cues can leave a customer confused or lost. Have you ever been unsure how to find your hotel room, even after the front-desk staff provided detailed directions? Better, clearer cues along the way would have enhanced your experience. Standard Parking of Chicago decorates each floor of its O'Hare Airport garage with icons of different Chicago sports franchises – the Bulls on one floor, the White Sox on another, and so forth. And each level has its own signature song wafting through it. "You never forget where you parked," one Chicago resident remarked, which is precisely the experience a traveler wants after returning from a week of travel.

Eliminate negative cues. Ensuring the integrity of the customer experience requires more than the layering on of positive cues. Experience stagers also must eliminate anything that diminishes, contradicts, or distracts from the theme. Most constructed

spaces – malls, offices, buildings, or airplanes – are littered with meaningless or trivial messages. While customers sometimes do need instructions, too often service providers choose an inappropriate medium or message form. For example, trash bins at fast-food facilities typically display a "Thank You" sign. True, it's a cue to customers to bus their own trays, but it also says, "No service here," a negative reminder. Experience stagers might, instead, turn the trash bin into a talking, garbage-eating character that announces its gratitude when the lid swings open. Customers get the same message but without the negative cue, and self-busing becomes a positive part of the eating experience.

The easiest way to turn a service into an experience is to provide poor service – thus creating a memorable encounter of the unpleasant kind. "Overservicing" in the name of customer intimacy can also ruin an experience. Airline pilots interrupt customers who are reading, talking, or napping to announce, "Toledo is off to the right side of the aircraft." At hotels, front-desk personnel interrupt face-to-face conversations with guests to field telephone calls. In the guestrooms, service reminders clutter end tables, dressers, and desktops. (Hide them away and housekeeping will replace these annoyances the next morning.) Eliminating negative cues – by transmitting pilots' offhand announcements through headsets instead of speakers, by assigning off-stage personnel to answer phones, and by placing guest information on an interactive television channel – creates a more pleasurable customer experience.

Mix in memorabilia. Certain goods have always been purchased primarily for the memories they convey. Vacationers buy postcards to evoke a treasured sight, golfers purchase a shirt or cap with an embroidered logo to recall a course or round, and teenagers obtain T-shirts to remember a rock concert. They purchase such memorabilia as a physical reminder of an experience.

People already spend tens of billions of dollars every year on memorabilia. These goods generally sell at price points far above those commanded by similar items that don't represent an experience. A Rolling Stones concert-goer, for example, will pay a premium for an official T-shirt emblazoned with the date and city of the concert. That's because the price points are a function less of the cost of goods than of the value the buyer attaches to remembering the experience.

If service businesses like airlines, banks, grocery stores, and insurance companies find no demand for memorabilia, it's because they do not stage engaging experiences. But if these businesses offered themed experiences layered with positive cues and devoid of negative cues, their guests would want and would pay for memorabilia to commemorate their experiences. (If guests didn't want to, it probably would mean the experience wasn't great.) The special agents of the Geek Squad, for example, stage such a distinctive computer-repair experience that customers buy T-shirts and lapel pins from the company's Web site. If airlines truly were in the experience-staging business, more passengers would actually shop in those seat-pocket catalogs for appropriate mementos. Likewise, mortgage loans would inspire household keepsakes; grocery checkout lanes would stock souvenirs in lieu of nickel-and-dime impulse items; and perhaps even insurance policy certificates would be considered suitable for framing.

If airlines truly sold experiences, more passengers would actually shop in the seat-pocket catalogs for mementos of their flight.

Engage all five senses. The sensory stimulants that accompany an experience should support and enhance its theme. The more senses an experience engages, the more effective and memorable it can be. Smart shoeshine operators augment the smell of polish with crisp snaps of the cloth, scents and sounds that don't make the shoes any shinier but do make the experience more engaging. Savvy hair stylists shampoo and apply lotions not simply because the styling requires it but because they add more tactile sensations to the customer experience. Similarly, grocery stores pipe bakery smells into the aisles, and some use light and sound to simulate thunderstorms when misting their produce.

The mist at the Rainforest Cafe appeals serially to all five senses. It is first apparent as a sound: Sss-sss-zzz. Then you see the mist rising from the rocks and feel it soft and cool against your skin. Finally, you smell its tropical essence, and you taste (or imagine that you do) its freshness. What you can't be is unaffected by the mist.

Some cues heighten an experience through a single sense affected through striking simplicity. The Cleveland Bicentennial Commission spent $4 mil-

lion to illuminate eight automobile and railroad bridges over the Cuyahoga River near a nightspot area called the Flats. No one pays a toll to view or even to cross these illuminated bridges, but the dramatically lighted structures are a prop that city managers hope will help attract tourist dollars by making a trip to downtown Cleveland a more memorable nighttime experience.

Not all sensations are good ones, and some combinations don't work. Bookstore designers have discovered that the aroma and taste of coffee go well with a freshly cracked book. But Duds n' Suds went bust attempting to combine a bar and a coin-operated laundromat. The smells of phosphates and hops, apparently, aren't mutually complementary.

Entering the Experience Economy

Using these five design principles, of course, is no guarantee of success; no one has repealed the laws of supply and demand. Companies that fail to provide consistently engaging experiences, overprice their experiences relative to the value perceived, or overbuild their capacity to stage them will of course see pressure on demand, pricing, or both. One stalwart of the children's birthday-party circuit, Discovery Zone, has had a rough few years because of inconsistent experience staging, poorly maintained games, and little consideration of the experience received by adults, who are, after all, paying for the event. More recently, the Rainforest Cafe and Planet Hollywood have encountered trouble because they have failed to refresh their experiences. Guests find nothing different from one visit to the next. Disney, on the other hand, avoids staleness by frequently adding new attractions and even whole parks such as the Animal Kingdom, which opened in the spring of 1998.

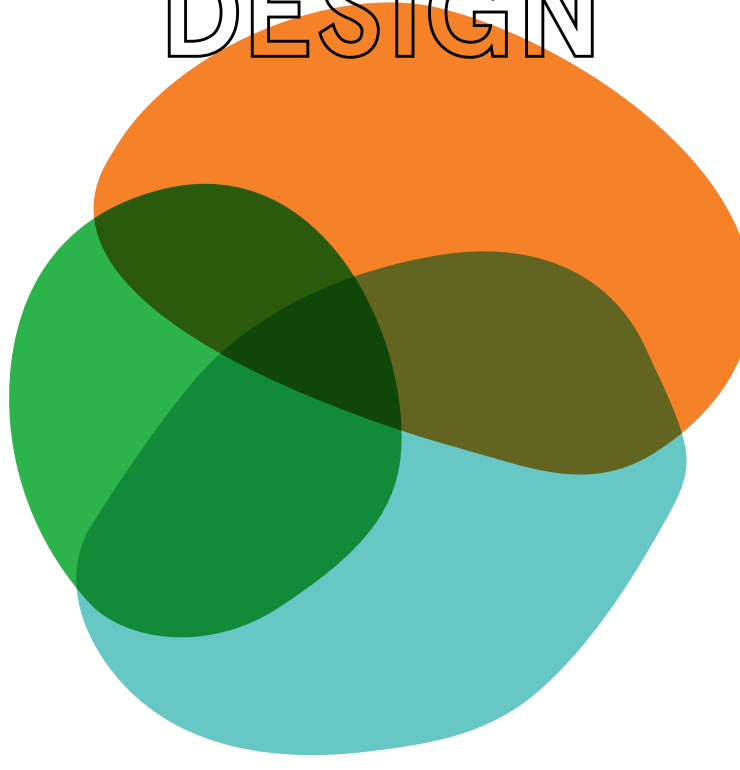
As the experience economy unfolds, more than a few experience stagers will exit the business. It's hard to imagine, for example, that every one of the scores of theme-based restaurants operating today will last into the millennium. Recall that once there were more than 100 automakers in eastern Michigan and more than 40 cereal makers in western Michigan. Now only the Big Three automakers in Detroit and the Kellogg Company in Battle Creek remain. The growth of the industrial economy and the service economy came with the proliferation of offerings – goods and services that didn't exist before imaginative designers and marketers invented and developed them. That's also how the experience economy will grow: through the "gales of creative destruction," as the economist Joseph Schumpeter termed it – that is, business innovation, which threatens to render irrelevant those who relegate themselves to the diminishing world of goods and services.

Reprint 98407 To place an order, call 1-800-988-0886.


POLICE


WA
~EASTHER
MILK BAR
SmartLife

THE FIELD GUIDE TO HUMAN-CENTERED DESIGN



DESIGN KIT

What Does It Mean to Be a Human-Centered Designer?

Embracing human-centered design means believing that all problems, even the seemingly intractable ones like poverty, gender equality, and clean water, are solvable. Moreover, it means believing that the people who face those problems every day are the ones who hold the key to their answer. Human-centered design offers problem solvers of any stripe a chance to design with communities, to deeply understand the people they're looking to serve, to dream up scores of ideas, and to create innovative new solutions rooted in people's actual needs.

At IDEO.org and IDEO, we've used human-centered design for decades to create products, services, experiences, and social enterprises that have been adopted and embraced because we've kept people's lives and desires at the core. The social sector is ripe for innovation, and we've seen time and again how our approach has the power to unlock real impact. Being a human-centered designer is about believing that as long as you stay grounded in what you've learned from people, your team can arrive at new solutions that the world needs. And with this Field Guide, you're now armed with the tools needed to bring that belief to life.

Adopt the Mindsets

Human-centered designers are unlike other problem solvers—we tinker and test, we fail early and often, and we spend a surprising amount of time not knowing the answer to the challenge at hand. And yet, we forge ahead. We're optimists and makers, experimenters and learners, we empathize and iterate, and we look for inspiration in unexpected places. We believe that a solution is out there and that by keeping focused on the people we're designing for and asking the right questions, we'll get there together. We dream up lots of ideas, some that work and some that don't. We make our ideas tangible so that we can test them, and then we refine them. In the end, our approach amounts to wild creativity, to a ceaseless push to innovate, and a confidence that leads us to solutions we'd never dreamed of when we started. In the Field Guide, we share our philosophy of design and the seven mindsets that set us apart: Empathy, Optimism, Iteration, Creative Confidence, Making, Embracing Ambiguity, and Learning from Failure.

Understand the Process

Human-centered design isn't a perfectly linear process, and each project invariably has its own contours and character. But no matter what kind of design challenge you've got, you'll move through three main phases: Inspiration, Ideation, and Implementation. By taking these three phases in turn, you'll build deep empathy with the communities and individuals you're designing for; you'll figure out how to turn what you've learned into a chance to design a new solution; and you'll build and test your ideas before finally putting them out into the world. At IDEO.org and IDEO, we've used human-centered design to tackle a vast array of design challenges, and though our projects have ranged from social enterprises to communication campaigns to medical devices, this particular approach to creative problem solving has seen us through each time.

INSPIRATION

In this phase, you'll learn how to better understand people. You'll observe their lives, hear their hopes and desires, and get smart on your challenge.

IDEATION

Here you'll make sense of everything that you've heard, generate tons of ideas, identify opportunities for design, and test and refine your solutions.

IMPLEMENTATION

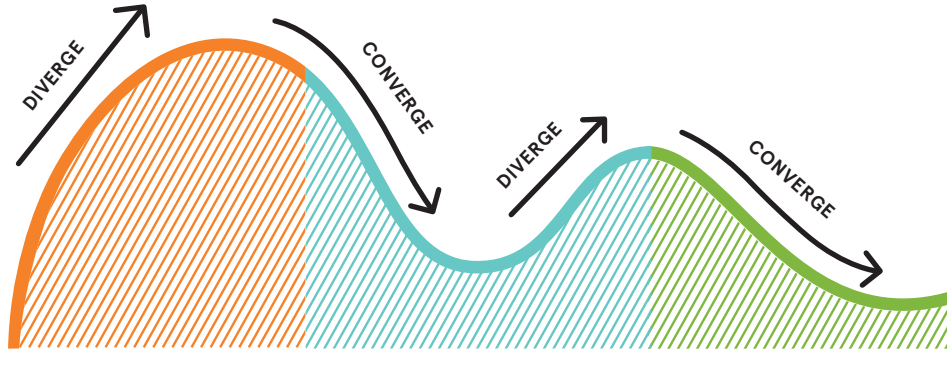
Now is your chance to bring your solution to life. You'll figure out how to get your idea to market and how to maximize its impact in the world.

Use the Tools

Though no two human-centered design projects are alike, we draw from the same kit of tools for each of them. For example, to build deep empathy with the people we're trying to serve, we always conduct interviews with them. To maintain creativity and energy, we always work in teams. To keep our thinking generative, sharp, and because it helps us work things through, we always make tangible prototypes of our ideas. And because we rarely get it right the first time, we always share what we've made, and iterate based on the feedback we get. The 57 methods in the Field Guide offer a comprehensive set of exercises and activities that will take you from framing up your design challenge to getting it to market. You'll use some of these methods twice or three times and some not at all as you work through your challenge. But taken as a set, they'll put you on the path to continuous innovation while keeping the community you're designing for squarely at the center of your work.

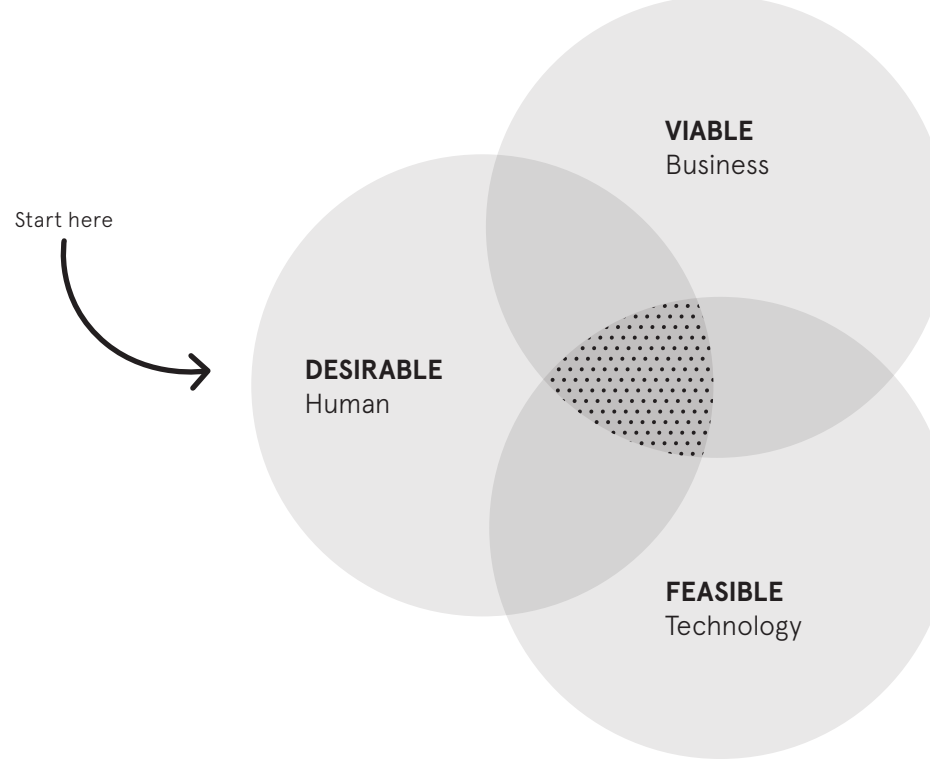
Trust the Process Even if It Feels Uncomfortable

Human-centered design is a unique approach to problem solving, one that can occasionally feel more like madness than method—but you rarely get to new and innovative solutions if you always know precisely where you're going. The process is designed to get you to learn directly from people, open yourself up to a breadth of creative possibilities, and then zero in on what's most desirable, feasible, and viable for the people you're designing for. You'll find yourself frequently shifting gears through the process, and as you work through its three phases you'll swiftly move from concrete observations to highly abstract thinking, and then right back again into the nuts and bolts of your prototype. We call it diverging and converging. By going really big and broad during the Ideation phase, we dream up all kinds of possible solutions. But because the goal is to have a big impact in the world, we have to then identify what, among that constellation of ideas, has the best shot at really working. You'll diverge and converge a few times, and with each new cycle you'll come closer and closer to a market-ready solution.



Create Real Impact

Human-centered design is uniquely situated to arrive at solutions that are desirable, feasible, and viable. By starting with humans, their hopes, fears, and needs, we quickly uncover what's most desirable. But that's only one lens through which we look at our solutions. Once we've determined a range of solutions that could appeal to the community we're looking to serve, we then start to home in on what is technically feasible to actually implement and how to make the solution financially viable. It's a balancing act, but one that's absolutely crucial to designing solutions that are successful and sustainable.



MINDSETS

Creative Confidence

Creative confidence is the notion that you have big ideas, and that you have the ability to act on them.

—David Kelley, Founder, IDEO

Anyone can approach the world like a designer. Often all it takes to unlock that potential as a dynamic problem solver is a bit of creative confidence. Creative confidence is the belief that everyone is creative, and that creativity isn't the capacity to draw or compose or sculpt, but a way of understanding the world.

Creative confidence is the quality that human-centered designers rely on when it comes to making leaps, trusting their intuition, and chasing solutions that they haven't totally figured out yet. It's the belief that you can and will come up with creative solutions to big problems and the confidence that all it takes is rolling up your sleeves and diving in. Creative confidence will drive you to make things, to test them out, to get it wrong, and to keep on rolling, secure in the knowledge that you'll get where you need to go and that you're bound to innovate along the way.

It can take time to build creative confidence, and part of getting there is trusting that the human-centered design process will show you how to bring a creative approach to whatever problem is at hand. As you start with small successes and then build to bigger ones, you'll see your creative confidence grow and before long you'll find yourself in the mindset that you are a wildly creative person.

Make It

You're taking risk out of the process by making something simple first. And you always learn lessons from it.

—Krista Donaldson, CEO, D-Rev

As human-centered designers, we make because we believe in the power of tangibility. And we know that making an idea real reveals so much that mere theory cannot. When the goal is to get impactful solutions out into the world, you can't live in abstractions. You have to make them real.

Human-centered designers are doers, tinkerers, crafters, and builders. We make using anything at our disposal, from cardboard and scissors to sophisticated digital tools. We build our ideas so that we can test them, and because actually making something reveals opportunities and complexities that we'd never have guessed were there. Making is also a fantastic way to think, and it helps bring into focus the feasibility of our designs. Moreover, making an idea real is an incredibly effective way to share it. And without candid, actionable feedback from people, we won't know how to push our ideas forward.

As you move through the human-centered design process, it doesn't matter what you make, the materials you use, or how beautiful the result is, the goal is always to convey an idea, share it, and learn how to make it better.

Best of all, you can prototype anything at any stage of the process from a service model to a uniform, from a storyboard to the financial details of your solution. As human-centered designers, we have a bias toward action, and that means getting ideas out of our heads and into the hands of the people we're looking to serve.

Learn from Failure

Don't think of it as failure, think of it as designing experiments through which you're going to learn.

—Tim Brown, CEO, IDEO

Failure is an incredibly powerful tool for learning. Designing experiments, prototypes, and interactions and testing them is at the heart of human-centered design. So is an understanding that not all of them are going to work. As we seek to solve big problems, we're bound to fail. But if we adopt the right mindset, we'll inevitably learn something from that failure.

Human-centered design starts from a place of not knowing what the solution to a given design challenge might be. Only by listening, thinking, building, and refining our way to an answer do we get something that will work for the people we're trying to serve. "Fail early to succeed sooner" is a common refrain around IDEO, and part of its power is the permission it gives to get something wrong. By refusing to take risks, some problem solvers actually close themselves off from a real chance to innovate.

Thomas Edison put it well when he said, "I have not failed. I've just found 10,000 ways that won't work." And for human-centered designers, sorting out what won't work is part of finding what will.

Failure is an inherent part of human-centered design because we rarely get it right on our first try. In fact, getting it right on the first try isn't the point at all. The point is to put something out into the world and then use it to keep learning, keep asking, and keep testing. When human-centered designers get it right, it's because they got it wrong first.

Empathy

In order to get to new solutions, you have to get to know different people, different scenarios, different places.

—Emi Kolawole, Editor-in-Residence, Stanford University d.school

Empathy is the capacity to step into other people's shoes, to understand their lives, and start to solve problems from their perspectives. Human-centered design is premised on empathy, on the idea that the people you're designing for are your roadmap to innovative solutions. All you have to do is empathize, understand them, and bring them along with you in the design process.

For too long, the international development community has designed solutions to the challenges of poverty without truly empathizing with and understanding the people it's looking to serve. But by putting ourselves in the shoes of the person we're designing for, human-centered designers can start to see the world, and all the opportunities to improve it, through a new and powerful lens.

Immersing yourself in another world not only opens you up to new creative possibilities, but it allows you to leave behind preconceived ideas and outmoded ways of thinking. Empathizing with the people you're designing for is the best route to truly grasping the context and complexities of their lives. But most importantly, it keeps the people you're designing for squarely grounded in the center of your work.

Embrace Ambiguity

We want to give ourselves the permission to explore lots of different possibilities so that the right answer can reveal itself.

—Patrice Martin, Co-Lead and Creative Director, IDEO.org

Human-centered designers always start from the place of not knowing the answer to the problem they're looking to solve. And in a culture that can be too focused on being the first one to the right answer, that's not a particularly comfortable place to be. But by starting at square one, we're forced to get out into the world and talk to the people we're looking to serve. We also get to open up creatively, to pursue lots of different ideas, and to arrive at unexpected solutions. By embracing that ambiguity, and by trusting that the human-centered design process will guide us toward an innovative answer, we actually give ourselves permission to be fantastically creative.

One of the qualities that sets human-centered designers apart is the belief that there will always be more ideas. We don't cling to ideas any longer than we have to because we know that we'll have more. Because human-centered design is such a generative process, and because we work so collaboratively, it's easy to discard bad ideas, hold onto pieces of the so-so ones, and eventually arrive at the good ones.

Though it may seem counterintuitive, the ambiguity of not knowing the answer actually sets up human-centered designers to innovate. If we knew the answer when we started, what could we possibly learn? How could we come up with creative solutions? Where would the people we're designing for guide us? Embracing ambiguity actually frees us to pursue an answer that we can't initially imagine, which puts us on the path to routine innovation and lasting impact.

Optimism

Optimism is the thing that drives you forward.

—John Bielenberg, Founder, Future Partners

We believe that design is inherently optimistic. To take on a big challenge, especially one as large and intractable as poverty, we have to believe that progress is even an option. If we didn't, we wouldn't even try. Optimism is the embrace of possibility, the idea that even if we don't know the answer, that it's out there and that we can find it.

In addition to driving us toward solutions, optimism makes us more creative, encourages us to push on when we hit dead ends, and helps all the stakeholders in a project gel. Approaching problems from the perspective that you'll get to a solution infuses the entire process with the energy and drive that you need to navigate the thorniest problems.

Human-centered designers are persistently focused on what could be, not the countless obstacles that may get in the way. Constraints are inevitable, and often they push designers toward unexpected solutions. But it's our core animating belief—that every problem is solvable—that shows just how deeply optimistic human-centered designers are.

Iterate, Iterate, Iterate

By iterating, we validate our ideas along the way because we're hearing from the people we're actually designing for.

—Gaby Brink, Founder, Tomorrow Partners

As human-centered designers, we adopt an iterative approach to solving problems because it makes feedback from the people we're designing for a critical part of how a solution evolves. By continually iterating, refining, and improving our work, we put ourselves in a place where we'll have more ideas, try a variety of approaches, unlock our creativity, and arrive more quickly at successful solutions.

Iteration keeps us nimble, responsive, and trains our focus on getting the idea and, after a few passes, every detail just right. If you aimed for perfection each time you built a prototype or shared an idea, you'd spend ages refining something whose validity was still in doubt. But by building, testing, and iterating, you can advance your idea without investing hours and resources until you're sure that it's the one.

At base, we iterate because we know that we won't get it right the first time. Or even the second. Iteration allows us the opportunity to explore, to get it wrong, to follow our hunches, but ultimately arrive at a solution that will be adopted and embraced. We iterate because it allows us to keep learning. Instead of hiding out in our workshops, betting that an idea, product, or service will be a hit, we quickly get out in the world and let the people we're designing for be our guides.

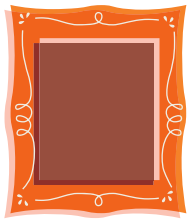
METHODS

INSPIRATION

The Inspiration phase is about learning on the fly, opening yourself up to creative possibilities, and trusting that as long as you remain grounded in desires of the communities you're engaging, your ideas will evolve into the right solutions. You'll build your team, get smart on your challenge, and talk to a staggering variety of people.

THIS PHASE WILL HELP YOU ANSWER

How do I get started?
How do I conduct an interview?
How do I keep people at the center of my research?
What are other tools I can use to understand people?



Frame Your Design Challenge

Properly framing your design challenge is critical to your success. Here's how to do it just right.

Getting the right frame on your design challenge will get you off on the right foot, organize how you think about your solution, and at moments of ambiguity, help clarify where you should push your design. Framing your design challenge is more art than science, but there are a few key things to keep in mind. First, ask yourself: Does my challenge drive toward ultimate impact, allow for a variety of solutions, and take into account context? Dial those in, and then refine it until it's the challenge you're excited to tackle.

TIME
90 minutes

DIFFICULTY
Hard

WHAT YOU'LL NEED
Pen, Frame Your Design Challenge worksheet p. 165

PARTICIPANTS
Design team

STEPS

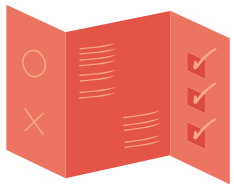
01 Start by taking a first stab at writing your design challenge. It should be short and easy to remember, a single sentence that conveys what you want to do. We often phrase these as questions which set you and your team up to be solution-oriented and to generate lots of ideas along the way.

02 Properly framed design challenges drive toward ultimate impact, allow for a variety of solutions, and take into account constraints and context. Now try articulating it again with those factors in mind.

03 Another common pitfall when scoping a design challenge is going either too narrow or too broad. A narrowly scoped challenge won't offer enough room to explore creative solutions. And a broadly scoped challenge won't give you any idea where to start.

04 Now that you've run your challenge through these filters, do it again. It may seem repetitive, but the right question is key to arriving at a good solution. A quick test we often run on a design challenge is to see if we can come up with five possible solutions in just a few minutes. If so, you're likely on the right track.

Create a Project Plan



Get organized, understand your strengths, and start identifying what your team will need to come up with innovative solutions.

As you set out to solve your challenge, you'll need to create a plan. This gives you a chance to think through all the logistics of your project, and even though they're bound to change as things progress, you'll be in much better shape if you can plan for what's ahead. Reflect on your timeline, the space you'll work in, your staff, your budget, what skills you'll need, trips you'll take, and what you'll need to produce. Getting a good handle on all of this information can keep you on track.

TIME
60-90 minutes

DIFFICULTY
Moderate

WHAT YOU'LL NEED
Pen, paper, Post-its, calendar

PARTICIPANTS
Design team

STEPS

01 A good place to start is with a calendar. Print out or make a large one and put it up in your workspace. Now mark key dates. They could be deadlines, important meetings, travel dates, or times when your team members are unavailable.

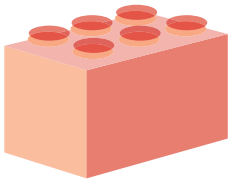
02 Now that you've got a sense of your timeline, look at your budget and staff. Do you have everything that you'll need? If you foresee constraints, how can you get around them?

03 You'll need to get smart on your topic before you head into the field. Who should you talk to now? What will you need to read to be up to speed on your challenge?

04 Answer questions like: When should my team head into the field? Will my team make one visit or two? Will our partners be visiting? Will we need to physically make something? How much time, money, and manpower will we need to produce it?

05 Your project plan will change as things evolve, and that's perfectly OK. You can always amend things as you go but make sure that you're really thinking through your project before you start.

Build a Team



An interdisciplinary mix of thinkers, makers, and doers is just the right combination to tackle any design challenge.

Human-centered design works best with cross-disciplinary teams. You could put three business designers to work on a new social enterprise, but if you throw a graphic designer, a journalist, or an industrial designer into the mix, you're going to bring new modes of thinking to your team. It's smart to have a hunch about what kind of talent your team will need—if you're designing a social enterprise, a business designer is probably a good bet—but you won't get unexpected solutions with an expected team.

TIME
60 minutes

DIFFICULTY
Hard

WHAT YOU'LL NEED
Pen, paper

PARTICIPANTS
Project lead, partner organizations

STEPS

01 First, assess how many team members you'll need, your staff's availability, and when your project should start and end.

02 Look at the core members of your team and determine what they're good at and what they're not so good at.

03 Is there a clear technical capability that you'll need but don't currently have—maybe a mechanical engineer, a graphic designer, a skilled writer? Remember that you can always add a team member for a shorter period of time when their skills are most important.

Recruiting Tools

Human-centered design isn't just about talking to a lot of people, it's about talking to the right people. Build a strategy now so that your Interviews really count.

Before you start talking to the people you're designing for, it's important to have a strategy around who you talk to, what you ask them, and what pieces of information you need to gather. By planning ahead, and tracking who you talk to once you've done it, you can be sure to have the right balance of experts and laymen, women and men, people of different ethnicities and classes, as well as a full range of behaviors, beliefs, and perspectives.

TIME
30-60 minutes

DIFFICULTY
Moderate

WHAT YOU'LL NEED
Pen, paper

PARTICIPANTS
Design team

STEPS

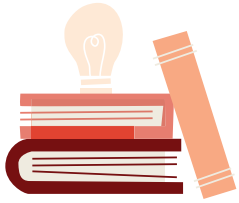
01 As you start to determine who you want to talk to, think about a variety of factors: age, gender, ethnicity, class, social position. Who do you really need to hear from?

02 Be sensitive to gender when making your Interview plan. Some communities may not be comfortable with men interviewing women. Or if you're working on a touchy topic, like open defecation, make sure that you understand social dynamics before you begin your Interviews (p. 39).

03 Group Interviews (p. 42) can be a highly useful tool and also help you identify who you might like to speak more with in an individual Interview.

04 Refer to Extremes and Mainstreams (p. 49) to make sure that you're talking to a broad spectrum of people.

Secondary Research



Getting up to speed on your challenge is crucial to success in the field.

Human-centered design is all about talking with people about their challenges, ambitions, and constraints. But as you move through the Inspiration phase, there will be moments where you'll need more context, history, or data than a man-on-the-street style Interview can afford. Social sector challenges can be really thorny, which is why Secondary Research, whether done online, by reading books, or by crunching numbers, can help you ask the right questions. At IDEO.org, we find that a firm foundation of knowledge is the best place from which to tackle a design challenge.

TIME
1-2 days

DIFFICULTY
Moderate

WHAT YOU'LL NEED
Internet connection, pen, notebook, research materials

PARTICIPANTS
Design team

STEPS

01 Once you know your design challenge, it's time to start learning about its broader context. You can bone up quickly by exploring the most recent news in the field. Use the Internet, newspapers, magazines, or journals to know what's new.

02 Try to find recent innovations in your particular area. They could be technological, behavioral, or cultural. Understanding the edge of what's possible will help you ask great questions.

03 Take a look at other solutions in your area. Which ones worked? Which ones didn't? Are there any that feel similar to what you might design? Any solutions that have inspired you to make one of your own?

04 Because Interviews (p. 39) can be highly subjective, use your Secondary Research to get the facts and figures you'll need to understand the context of your challenge.

Interview

There's no better way to understand the hopes, desires, and aspirations of those you're designing for than by talking with them directly.

Interviews really are the crux of the Inspiration phase. Human-centered design is about getting to the people you're designing for and hearing from them in their own words. Interviews can be a bit daunting, but by following these steps below you'll unlock all kinds of insights and understanding that you'll never get sitting behind your desk. Whenever possible, conduct your Interviews in the person's space. You can learn so much about a person's mindset, behavior, and lifestyle by talking with them where they live or work.

TIME
60-90 minutes

DIFFICULTY
Moderate

WHAT YOU'LL NEED
Pens, paper, Interview Guide worksheet p. 166

PARTICIPANTS
Design team, person you're designing for

STEPS

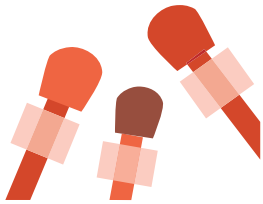
01 No more than three research team members should attend any single Interview so as to not overwhelm the participant or crowd the location. Each team member should have a clear role (i.e. interviewer, note-taker, photographer).

02 Come prepared with a set of questions you'd like to ask. Start by asking broad questions about the person's life, values, and habits, before asking more specific questions that relate directly to your challenge.

03 Make sure to write down exactly what the person says, not what you think they might mean. This process is premised on hearing exactly what people are saying. If you're relying on a translator, make sure he or she understands that you want direct quotes, not the gist of what the person says.

04 What you hear is only one data point. Be sure to observe the person's body language and surroundings and see what you can learn from the context in which you're talking. Take pictures, provided you get permission first.

Group Interview



You can come to a quick understanding of a community's life, dynamics, and needs by conducting a Group Interview.

Though a Group Interview may not offer the depth of an individual Interview (p. 39) in someone's home, it can give you a compelling look at how a larger set of the people you're designing for operates. The best Group Interviews seek to hear everyone's voice, get diverse opinions, and are strategic about group makeup. For example, an all-female group might give you insight into the role of women in a society whereas a mixed group may not. If you're looking to learn quickly what is valuable to a community, Group Interviews are a great place to start.

TIME
90-120 minutes

DIFFICULTY
Moderate

WHAT YOU'LL NEED
Pens, paper, camera

PARTICIPANTS
At least 2 members of the design team, 7-10 people you're designing for

STEPS

01 Identify the sort of group you want to talk with. If you're trying to learn something specific, organize the group so that they're likely to have good answers to the questions that you've got.

02 Convene the Group Interview on neutral ground, perhaps a shared community space that people of all ages, races, and genders can access.

03 In a Group Interview, be certain to have one person asking the questions and other team members taking notes and capturing what the group is saying.

04 Come prepared with a strategy to engage the quieter members of the group. This can mean asking them questions directly or finding ways to make the more vocal members of the group recede for a moment.

05 Group Interviews are a great setting to identify who you might want to go deeper with in a Co-Creation Session (p. 109).

Expert Interview

Experts can fill you in quickly on a topic, and give you key insights into relevant history, context, and innovations.

Though the crux of the Inspiration phase is talking with the people you're designing for, you can gain valuable perspective by talking to experts. Experts can often give you a systems-level view of your project area, tell you about recent innovations—successes and failures—and offer the perspectives of organizations like banks, governments, or NGOs. You might also look to experts for specific technical advice.

TIME
60-90 minutes

DIFFICULTY
Moderate

WHAT YOU'LL NEED
Pens, camera, notebook

PARTICIPANTS
Design team, expert

STEPS

01 Determine what kind of expert you need. If you're working in agriculture, perhaps an agronomist. In reproductive health? A doctor or policymaker may be a good bet.

02 When recruiting your experts, give them a preview of the kinds of questions you'll be asking and let them know how much of their time you'll need.

03 Choose experts with varying points of view. You don't want the same opinions over and over.

04 Ask smart, researched questions. Though you should come prepared with an idea of what you'd like to learn, make sure your game plan is flexible enough to allow you to pursue unexpected lines of inquiry.

05 Record your Expert Interview with whatever tools you have. A pen and paper work fine.

Define Your Audience

Consider the broad spectrum of people who will be touched by your design solution.

Before you dig into your in-context research, it's critical to know who you're designing for. You're bound to learn more once you're in the field, but having an idea of your target audience's needs, contexts, and history will help ensure that you start your research by asking smart questions. And don't limit your thinking just to the people you're designing for. You may need to consider governments, NGOs, other businesses, or competitors.

TIME
30-60 minutes

DIFFICULTY
Easy

WHAT YOU'LL NEED
Pen, paper, Post-its

PARTICIPANTS
Design team

STEPS

01 With your team, write down the people or groups that are directly involved in or reached by your project. Are you designing for children? For farmers? Write all the groups down on Post-its and put them on a wall so you can visualize your audience.

02 Now add people or groups who are peripherally relevant, or are associated with your direct audience.

03 Think about the connections these people have with your topic. Who are the fans? Who are the skeptics? Who do you most need on your side? Add them to the wall.

04 Now arrange these Post-its into a map of the people involved in your challenge. Save it and refer to it as you move through the Inspiration phase.

CHAPTER FOUR

Who's Afraid of Machine Learning?

Who's afraid of a little machine learning (ML)? To understand this technology, designers need to get into the weeds—just a little. There are three main types of machine-learning algorithms: supervised learning, unsupervised learning, and reinforcement learning. Let's start with supervised learning.

SUPERVISED LEARNING

Supervised learning relies upon a full set of labeled data. (Labeled data means data that has been tagged, effectively placing the data into categories. Think spreadsheets of information.) Algorithms examine this labeled data, learn from patterns in the data, and then make predictions. It's the most prevalent form of machine learning today; therefore, it's worth examining closely.

Researchers would choose supervised learning to take on problems of *classification* or problems of *regression*. If we want the algorithms to predict the discrete category or "class" that new data will fall within, we would use classification. If we want the algorithms to predict outputs related to a real-valued number, we would use *regression*.

If you are feeling overwhelmed with technical jargon right now, bear with me. I'll explain each of these.

CORN SNAKE OR COPPERHEAD?

IF IT'S IN YOUR YARD, YOU NEED TO KNOW.

Let's start with a supervised learning strategy that uses classification. Imagine that snakes begin to infest your neighborhood. Neighbors begin posting pictures online to determine whether the particular snake in their own yard is a copperhead—poisonous—or the similar-looking corn snake—harmless. Their kids play outside, and they need to know. You decide to design an app to solve this problem. To develop the app, you need a system for differentiating snake species.

First, you gather training data: a set of snakes individually labeled as either a copperhead or a corn snake. This data composes our "ground truth." We can also refer to these two labels as classes.

Next, we need to determine features that the system can use to distinguish one class from the other. In this example, let's identify two features: length and mass. (Note that this is only for the sake of our example. Do not actually use these features to classify snakes and risk the life of a kid.)

A data scientist, developer, or designer might determine these specific features, or a team of experts in a particular domain might select the features and label the data. This is called annotation. In our example, we might bring in a team of expert herpetologists. The herpetologists would create a document with a set of rules for annotation. This process can be straightforward, but it can also be quite controversial. Participants might disagree over which features to use or what dataset is most appropriate. Domain-specific experts

LABELED DATA

SNAKE	LENGTH	MASS
Copperhead	11.7	3.4
Copperhead	10.2	4.4
Corn Snake	12.8	5.0
Copperhead	12.0	3.0
Corn Snake	12.4	4.8
Corn Snake	11.9	4.5

might also be worried about sharing confidential knowledge about the data with the data science team running the ML algorithms.

Let's assume that our process goes smoothly. Once the features are selected, we clean the resulting labeled data, ridding it of errors, inconsistencies, missing data, and duplicates. We then randomize the order of the data and look for bias in the training examples. If there are too many copperheads used in the training data, more copperheads will subsequently be identified in the system, biasing it toward copperheads. We've seen this play out in numerous real-world examples when, for instance, an ML system has been trained using too many examples of one skin tone and not enough of others.[1]

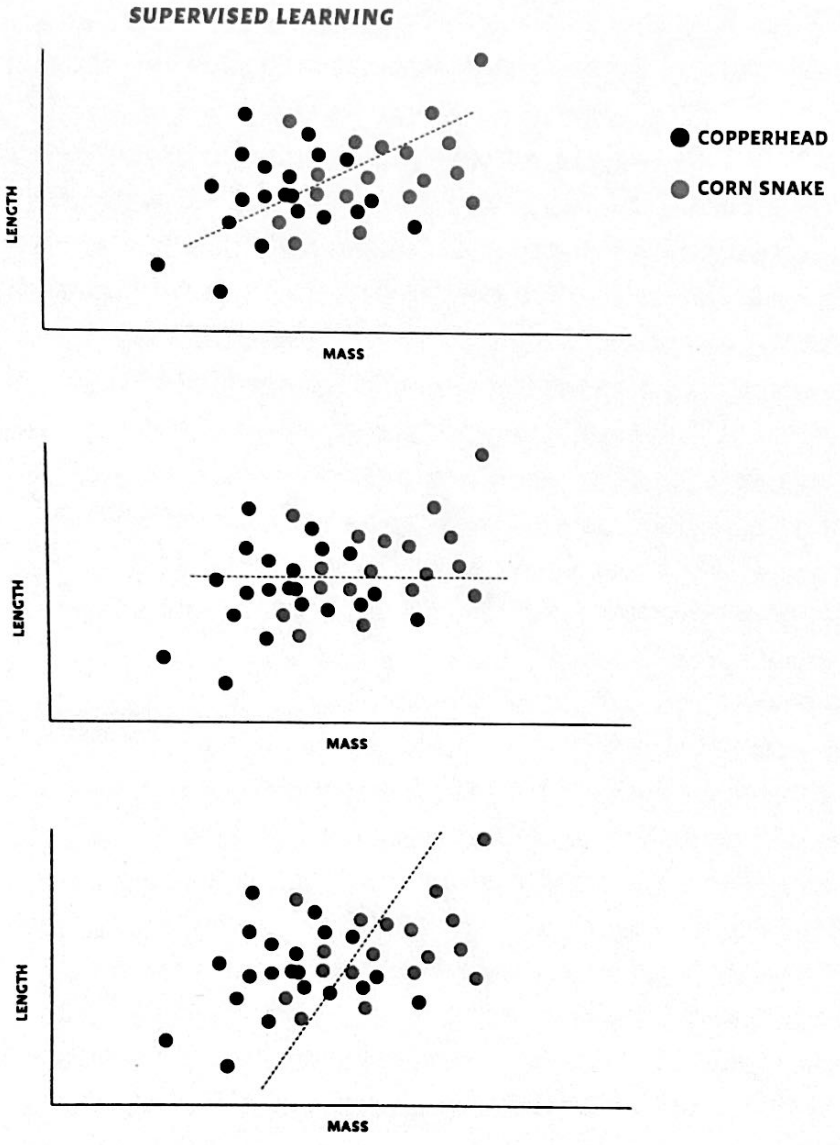
We then set aside a portion of the data to use to evaluate our system later.

STATISTICAL MODEL: a mathematical representation of data based on relationships observed within the data

Now the fun begins. We choose a learning algorithm and instruct that algorithm to build a model from the training set. In this example, we need to teach the system to separate the data into two different categories—or classes—of snake species so utilizing a classification model makes sense.

Note that this approach, like all ML, builds a statistical model. We can't achieve 100 percent accuracy with any model, but we will work to select and tweak our model to be as accurate as possible. ML, like design, is iterative. To increase the predictive power, we will need to refine the model throughout the process.

There are multiple strategies that a classification model could employ to achieve our goal, such as Decision Tree, K-Nearest Neighbor, and Naive Bayes. In this instance, let's select a Decision Tree strategy. When we run this selected strategy, the learning algorithms will, in essence, create a model that draws a line through the data creating a boundary separation. Based upon their features, some of the snakes will fall under the class of copperhead and some will fall under the class of corn snake. The algorithms will then check the results against the training data for accuracy and redraw the line



repeatedly to find the optimal boundary separation—the position in which the most snakes are accurately classified under their respective snake species.

Humans oversee this process, working to improve the results. As noted previously, the training data matters. We may need to add or delete training examples, particularly outliers, to increase accuracy. Or we may need to adjust the features. We may even try a different classification strategy, like K-Nearest-Neighbor or Naive Bayes. Each strategy will have its own advantages and trade-offs.

OUTLIERS: atypical values within a dataset; values that lie outside the distribution pattern

When we feel confident about the results, we can run the model on the labeled data that we set aside earlier for the evaluation phase. Depending on the results, we can keep working on the model or accept the accuracy rate.

Eventually, we should be able to run any copperhead or corn snake through this system and identify its species at a high rate of accuracy, not 100 percent but a high rate. We call the computer's decision that a certain snake is a copperhead or corn snake a prediction.

Let's step back and appreciate the beauty of the resulting system. We no longer need human judgment to determine the species of snake. And we don't need to program manual rules. In addition, if we have built-in feedback, the algorithms can learn from each misclassification and continually improve the results once the system is in play. Phew, the kids—and the corn snakes—are safe!

In this example we used two classes—copperhead and corn snake—and two features—mass and length. Keep in mind that predictive algorithms can use many, many classes and hundreds if not thousands of features.

REGRESSION

Regression differs from classification because it allows us to explore values—and thus make predictions—in between and beyond discrete classes. Regression can do this because the output is numerical (or continuous). We wouldn't pick a regression model for classifying snakes because, in our example, we want each snake to be identified as either one species or another.

We would use regression in situations in which we want the system to identify values in between or beyond what we initially defined in the training data. Such values might include the future price of something, customer-satisfaction level, or the grades of a student. In each of these instances, the algorithms would use the relationship between the continuous number and some other variable(s) to make predictions.

We used a classification strategy to organize input data (our snakes) into discrete classes. Using regression, we want the algorithms to predict a specific numerical value rather than a class. Models commonly used in regression include linear models, polynomial regression models, and, for more complex regression problems, neural networks.

For example, we might use regression to predict the future price of a house depending upon the rate of local job growth, or the level of customer satisfaction depending upon wait time, or the grade of a student in relation to the hours spent working in a studio.

Let's envision the process of running a model to predict a student's grade (a numerical value between 0 and 100). First we would label training examples of student grades, the dependent variable, paired with an independent variable, such as the amount of time physically spent weekly in the studio. Once we established some labeled examples, we could train the ML algorithms to predict the future grade of any student based on the independent variable—the amount of studio time. Now, think bigger. Just as with classification, predictive algorithms could compute a problem much more complex than our simple example. Instead of one variable, in this case time spent in the studio, there could be many.

DEPENDENT VARIABLE: a variable whose value depends on that of another

INDEPENDENT VARIABLE: the input for a process that is being analyzed (i.e., the feature)

Note that even though with regression the algorithms can make value predictions beyond those explicitly expressed in the training data (i.e., we don't have to provide examples of every possible grade/hour combination), they are doing so based on labeled data and defined variables. When we get to unsupervised learning, this will change.

SUPERVISED LEARNING: KEY TAKEAWAYS

- requires labeled training data
- has a clearly defined goal: "Computer, look for these particular patterns in the data so that you might predict x."
- the most prevalent form of ML today

UNSUPERVISED LEARNING

In unsupervised learning, an expert does not label the training data or provide the features/variables. Instead, the algorithms parse through input data looking for regularities or patterns that have not been prespecified. So, for example, instead of feeding our algorithms lots of images of snakes, already labeled as corn snake or copperhead, we might just give the system lots of snake images of both species and ask the algorithms to look for patterns to differentiate between the two. The algorithms themselves would then identify variables to reveal patterns. We wouldn't necessarily know how the algorithms selected the variables—and there could be thousands. We could even ask the algorithms to look for interesting patterns with less of a prescribed goal in mind.

Supervised learning requires a supervisor—someone to label the data. Unsupervised learning doesn't. Bypassing the supervisor can result in striking, unexpected outcomes.

When might humans use unsupervised learning? We might select unsupervised learning in situations in which the outcome is unclear or the analysis process too complex for a human to determine key differentiating labels and variables. Such complex situations might include identifying a human face, fraud detection, or predicting consumer behavior. Unsupervised learning also plays a key role in deep learning—to be discussed later.

When ML algorithms detect complex patterns using thousands of variables, humans sometimes struggle to understand how the machines arrived at their predictions. We call this the Black Box Problem. This might not be a big deal when a targeted ad miscalculates our interests. However, if we were denied parole based on a machine prediction of recidivism, it would suddenly be a huge deal—particularly if no one could explain how the decision was made.[2]

In addition to successfully analyzing highly complex situations, unsupervised learning algorithms can cut down on the time and expense necessitated by labeled data. Remember that

supervised learning algorithms require labeled data and identified features for training. Cheaper, more easily acquired, unlabeled data can be fed into unsupervised learning algorithms. All that unstructured, multimodal data—images, sounds, movement—referenced earlier becomes rich fodder for these systems.[3]

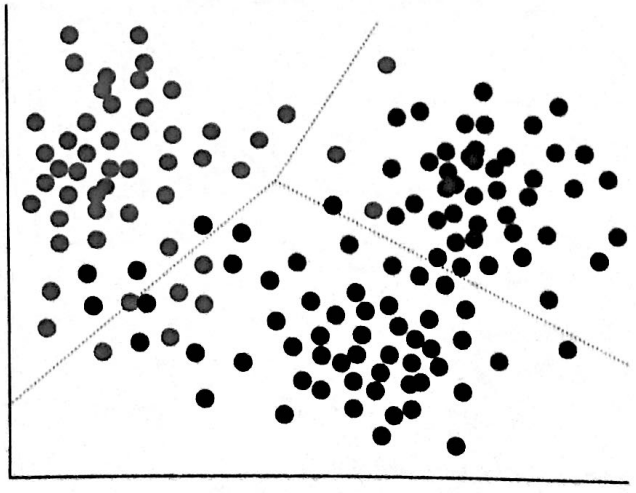
Common strategies in unsupervised learning include clustering and dimensionality reduction. Let's look at clustering.

CLUSTERING

Using this strategy, algorithms compute clusters or groups within the input data, i.e., data points with similar features are grouped together and data points with dissimilar features apart. Popular clustering algorithms include K-Means, Mean Shift, DBSCAN, Expectation-Maximation Clustering using Gaussian Mixture Models, and Agglomerative Hierarchical Clustering.

A retail company might use clustering to segment their customers. They could provide existing customer data points as input—age, zip code, gender, purchase history, etc.—and then run clustering algorithms to establish new groupings. The algorithms might discover customer segments that a designer or marketer would not typically envision. Or the model might reveal outliers that could inspire niche markets.[4] We could

UNSUPERVISED LEARNING: CLUSTERING



also use clustering to identify future trends like the professional turnover rate in a particular industry.[5] Darktrace, a cybersecurity company, uses unsupervised learning to learn patterns of normal behavior in a system, and then looks for anomalies—suspicious behavior—that could be cyberattacks. Rather than rely upon labeled data built from accumulated knowledge of past security threats, the system can look for threats that companies haven't yet identified.[6] Clustering can reveal unexpected insights that push beyond the typical human perspective.

UNSUPERVISED LEARNING: KEY TAKEAWAYS

- can detect patterns in situations too complex for human analysis
- does not require labeled data and/or identified features/variables
- can arrive at unexpected insights

AGENT: the one that decides what action to take in response to rewards or punishments

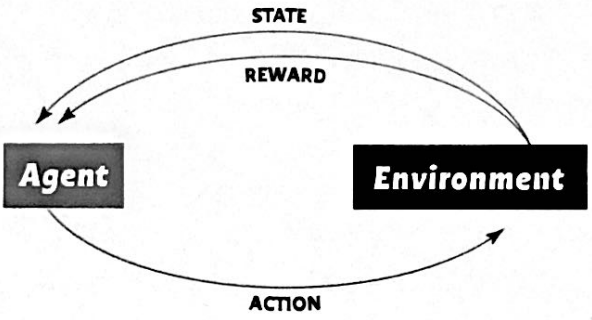
ENVIRONMENT: the surroundings or conditions within which an agent takes action

REINFORCEMENT LEARNING

In both supervised and unsupervised learning, algorithms make predictions based on training data. We can think of this as historical data because it already exists in the world. Knowledge of the past dictates predictions of the future. We talked about some of the negative implications of this in chapter three. Here, let's just consider the following, "Would you want to always be judged by your past behavior?"

In contrast, reinforcement learning algorithms do not make predictions based on historical data. Instead, these algorithms build a prediction model on the fly by interacting with an environment using trial and error. Here's how it works: an agent tries out possible actions within an environment. Each interaction with the environment produces insight—which acts as the input data. The agent then uses this data to iteratively adjust their actions to achieve a specific goal—a reward.

COMPONENTS OF REINFORCEMENT LEARNING



Think of a video game as an easy analogy for reinforcement learning. In such a game, players move through the game environment. By "playing," or interacting within that environment, they learn which actions move them toward the goal of winning and which do not. They repeat actions that work, using the resulting insight to inform subsequent behavior.

DEEP CONVOLUTIONAL NETWORK: a specific kind of deep neural network that has a convolution layer

This form of ML is most akin to the way we humans might acquire knowledge. To learn to roller skate, we put on skates and try. Each time we fall, we learn something about staying upright. We achieve our goal—moving around on skates while maintaining our balance—by putting this new knowledge into action.

In 2017, DeepMind employed reinforcement learning combined with deep convolutional networks, using a program called AlphaGo, to beat a human champion at the nuanced game of Go. Later the same year, they released a more advanced program called AlphaZero that mastered chess, shogi, and Go. According to DeepMind, the style and complexity of each game determined the training time: nine hours for chess, twelve hours for shogi, and thirteen days for Go.[7] Compare this to the years of intensive training that mastery requires of a human. Now imagine this technology applied to larger high-impact problems that require strategic action—global warming, for instance. Prior to this moment in 2017, according to Pedro Domingos, author of *The Master Algorithm*, "The supervised-learning people would make fun of the reinforcement-learning people."[8] DeepMind set everyone straight.

Because reinforcement learning algorithms can interact with an unpredictable digital or physical environment, we often see them employed in gaming AIs, logistics, resource management, robotics, and autonomous car navigation systems. Researchers can conveniently run these algorithms through millions of virtual simulations before taking their products into the world.[9] Computer scientist Mark Crowley, of the University of Waterloo, currently trains virtual fires, using reinforcement learning so that he might predict the path of future wildfires.[10] Uber refines their self-driving AI platform by putting it through thousands of virtual simulations, using predictive algorithms to play the self-driving AI against an equally intelligent environment.[11] The training period of these algorithms can be lengthy—typically longer than that of supervised learning systems despite the breakneck speed of AlphaZero—but the ability to respond to environments *in real time* can trump that substantial training time.[12]

Common reinforcement learning strategies include Deep Deterministic Policy Gradient, Q-Learning, State-Action-Reward-State-Action, and Deep Q-Networks. Note that in each of these strategies, as in unsupervised learning, no supervisor oversees the process. Because there is no supervisor, the supervisor cannot introduce bias into the system. To be clear, this does not eliminate all possibilities for bias, but it does get rid of one common avenue.[13] This lack of a supervisor also means that reinforcement learning can introduce alien strategies for "winning," i.e., achieving the reward, and then, in turn, teach these unusual strategies back to humans. In the previous example of AlphaGo and AlphaZero, Go champions studied the winning strategies employed by these programs and copied the tactics in their own subsequent games. Some of the tactics, however, were unusable because, try as they might, human players found them incomprehensible.[14] With less overt bias and the introduction of mind-boggling new strategies, reinforcement learning is pulling in big research dollars right now.[15]

REINFORCEMENT LEARNING: KEY TAKEAWAYS

- gains insight through trial and error
- model interacts with an environment sequentially over time
- eliminates supervisor bias
- can introduce alien approaches that prove useful to humans

CONCLUSION

Although we considered each ML category separately, many researchers mix these categories together in practice. Semi-supervised learning refers specifically to blends of supervised and unsupervised learning, but researchers can also use all three categories to achieve their goals.[16] This process can be messy and complex, requiring a deep knowledge of mathematics and statistics and a big helping of intuition. Training a set of algorithms requires a different skill set than programming explicit logic-based instructions.

SEMI-SUPERVISED LEARNING: a training approach in machine learning that combines supervised and unsupervised methods

This shift from programming to training produces a relationship with machines that is less clear cut and more difficult to control. Jason Tanz, the site director of *Wired*, explains, "If in the old view programmers were like gods, authoring the laws that govern computer systems, now they're like parents or dog trainers. And as any parent or dog owner can tell you, that is a much more mysterious relationship to find yourself in."[17] Designers thrive in this kind of liminal space. Working with data scientists, we can begin prototyping user experience and interface possibilities that map out such emerging relationships between human and machine.